

Deliverable 4.3

Map of Machinery

Due date of deliverable: 30/11/2024

Actual submission date: 30/11/2024

Start date of project: 01/02/2022 Duration (36 Months)

Dissemination Level: Public ✓



DELIVERABLE

Work Package	WP4 – Distributed Manufacturing Network
Deliverable	D4.3 Machinery – Map (M34) Online map of machinery based on Open Know-Where standard (min. 500 entries, 100 from the EU DIH networks) with forms (taken from TRL7 to TRL 9) and promotional material.
Document Name	Map of Machinery
Due Date	M34: 30 November 2024
Submission Date	M34: 30 November 2024
Dissemination Level	<input checked="" type="checkbox"/> P – Public <input type="checkbox"/> CO – Confidential
Deliverable Lead	Internet of Production (IOP)
Author(s)	Antonio Anaya; antonio.anaya@internetofproduction.org
Contributor(s)	Jessica Nguema; jessica.nguema@internetofproduction.org , Andrew Lamb; andrew.lamb@internetofproduction.org
Point of Contact	Barbara Kieslinger (kieslinger@zsi.at)
Reviewers	Fabienne Kirchhof; f.kirchhof@greentec-capital.com , Samuel Boadu Duah, samuel.b.duah@gmail.com
Status	<input type="checkbox"/> Plan <input type="checkbox"/> Draft <input type="checkbox"/> Working <input checked="" type="checkbox"/> Final <input type="checkbox"/> Approved
Abstract (for public dissemination only)	This document describes the development of online maps showing manufacturing capabilities and the usefulness of the Open Know-Where standard to share data on machines, facilities, people, and materials enabling distributed manufacturing.
Keywords	Map, Data processing, Use cases, Manufacturing resources



The information, documentation and figures in this deliverable are written by the mAKE project consortium under EC grant agreement number 101016858 and do not necessarily reflect the views of the European Commission. The European Commission is not liable for any use that may be made of the information contained herein.

Document History

Version	Date	Comment
001	13 09 2024	First structure and outline
002	25 11 2024	Main content added to all sections
003	26 11 2024	Peer review and detailed feedback on overall structure and specific content
004	27 11 2024	Integration of peer review comments
1.00	28 11 2024	Final editing

List of Tables

Table 1. Data source tables integrated on the Map of Machinery ETL system	20
Table 2. Technology Readiness Levels from 7 to 9	58

List of Abbreviations

API	Application Programming Interface
DIH	Digital Innovation Hub
DIY	Do-It-Yourself
DMP	Data Management Plan
DoA	Description of Action



DOI	Digital Object Identifier
DSI	Digital Social Innovation
EGE	The European Group on Ethics in Science and New Technologies
F.A.I.R.	Findable, Accessible, Interoperable, and Reusable
ETL	Extract, Transform, Load
GDPR	General Data Protection Regulations
GIS	Geographic Information System
HOT	Humanitarian OpenStreetMap Team
IMA	Innovative Manufacturing in Africa
IPR	Intellectual Property Rights
JSON	JavaScript Object Notation
NGO	Non-governmental organization
MVP	Minimum Viable Product
ODbL	Open Database License
OKH	Open Know-How
OKW	Open Know-Where
RISA	Research and Innovation Systems for Africa
UI	User Interface
STEM	Science, Technology, Engineering, and Mathematics
TRL	Technical Readiness Level



TABLE OF CONTENT

Executive Summary	7
Glossary and definitions	10
Introduction	15
1. The map of machinery	17
1.1. What is the Map of Machinery?	17
1.2. What are the use cases proposed for the map of machinery?	17
2. Data assessment	18
2.1. Details of data sources	23
FabLab Network, Fab Foundation, and partners	23
Maker Tour, and partners	24
Field Ready Inc, NeedsList, and partners	25
Makezine, makerspace.com, and partners	26
Hackerspaces.org, and partners	27
Verbund Offener-Werkstätten, and partners	28
Makery, and partners	29
Make Works, and partners	30
3. Proof of concept for data collection	31
3.1. Catalogue of machinery	32
4. Data specification	33
4.1. Data architecture	33
4.2. Data privacy policy	36
4.3. ETL System	37
5. Map design	42
6. Minimal Viable Product: A map for disaster response	44
7. Data collection Platform	47
7.1. Surveying tools	47
7.2. Data collection form design	48
7.3. Evaluate data collection process	51
8. Data and map usage	53
8.1. Online Maps	53
8.2. Data usage example	55
8.3. Data map analytics:	55
9. Publication & Promotion	57
10. Learnings	58



11. Technology Readiness Levels	59
12. References	60



Executive Summary

This document outlines the "Map of Machinery" project, deliverable 4.3 of the "mAKE African European Maker Innovation Ecosystem", commissioned and developed by the Internet of Production (IOP). This document details the motivations and processes behind the initial proposal, along with challenges, redefinitions, iterations, and outcomes, including data systems, visualizations, media resources, and software tools.

The overall objective is defined as a collection of information on locations, capabilities, and ownership of manufacturing machinery and displaying it in the form of an interactive geographical map. The project was planned and developed in accordance with the requirements, technical assessments, data privacy regulations, challenges, community input, and recommendations from the consortium board.

The value proposition of the Map of Machinery is to develop an accessible information system that fosters intelligence and collaboration within a distributed manufacturing network, connecting African makers with their counterparts in Europe. Mapping the geographical distribution of manufacturing capabilities across Africa provides tools and resources for governments to support innovation and for local and international organizations to engage small manufacturers in local procurement for humanitarian aid, disaster relief, education, and economic development. Additionally, it enables makers to explore, collaborate, and locate nearby and remote resources.

The project prioritizes open, replicable data aligned with F.A.I.R. principles, ensuring data veracity and quality assurance. Data collection presents technical challenges, as it requires a level of trust and motivation from collectors to ensure usability. In line with data governance best practices, the project conducted an account of online data already organized and collected by mAKE consortium partner organizations. This data retrieval and analysis helped define relevant fields, data volume, and purpose.

Through a series of documented interviews and working group meetings, a selection of attributes and motivations for data collection and mapping was synthesized into four primary use cases: 1) Finding Manufacturing Sites for Open-Source Projects, 2) Employment Creation Guidelines, 3) Maker-Space Learning and Setup Guidelines, and 4) Rapid Response Procurement Tool.

Case 4, Rapid Response Procurement, was selected for tool development with the mission to encourage data collection and promote the purpose of data sharing between partner organisations. The project produced a data policy document, database specifications, analysis reports, and computer code for the retrieval, wrangling and validation of data, as well as a self-hosted AI model for natural-language processing. Additionally, portable maps were developed for online and offline use, along with offline-first survey forms using KoBoToolbox. All tools adhere to open-source and compatible licensing to support the mission of creating an open information and data system.

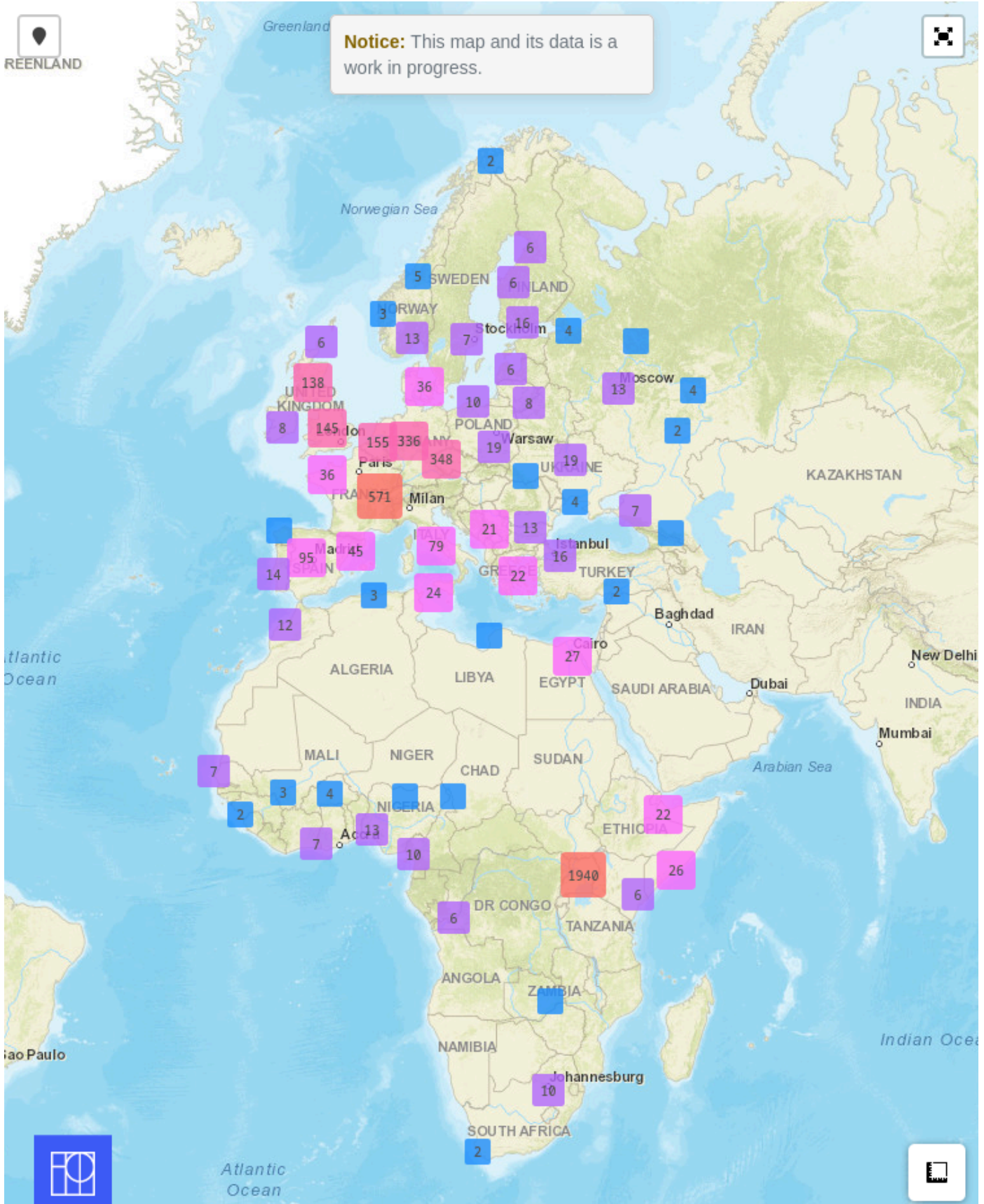


The state of collaborations between Distributed Manufacturing Networks was unknown, and unmapped for the African and European continents. At the time of writing the initial proposal, few data sources could tell the distribution of spaces and any other information of their relationships. A tool for finding and mapping manufacturing capabilities was hypothesised to serve for organising a perceived unstandardised global maker movement.

In 2020, the Internet of Production Alliance released the first version of the metadata specification of the Open Know-Where standard. "This standard has the mission to annotate, classify and collect data of manufacturing facilities, capabilities, and machines, and know-where to produce physical products following the objectives of Distributed manufacturing. The objective of an open standard for locating machines is simply to add machines to the map." (Lamb & Schack, 2021). In the Distributed Design book "Viral Design" it was mentioned, in regards to scaling in production, that "maker spaces and on-demand manufacturing can not only pursue production efficiencies solely for profit but can also serve the local needs of people. An inventory of local materials and manufacturing capacity could, for example, serve as a tool for designers, manufacturers and others that play an important role to revitalize local economies, especially after the pandemic." (Diez & Baeck, 2020).

The Map of machinery, based on the Open Know-Where standard, is an open innovative infrastructure enabling to find nearby the tools necessary for local production.





This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 101016858.

Glossary and definitions

Geographic Information System (GIS)

“Geographic information system(s), GIS (noun); GIS is a technology that is used to create, manage, analyse, and map all types of data. GIS connects data to a map, integrating location data (where things are) with all types of descriptive information (what things are like there). This provides a foundation for mapping and analysis that is used in science and almost every industry. GIS helps users understand patterns, relationships, and geographic context. The benefits include improved communication, efficiency, management, and decision-making.”

Retrieved from <https://www.esri.com/en-us/what-is-gis/overview>

Geography

“Geography is the scientific study of the Earth's physical features, climate, and resources, as well as the complex relationships between places, environments, and human societies. It encompasses the investigation of natural processes, human cultures, spatial patterns, and the ways in which societies interact with and organize space.”

Cresswell, T. (2013). "Geographic Thought: A Critical Introduction." Wiley-Blackwell.

Data

“Data refers to the raw, unorganized facts that need to be processed. Data can be something simple and seemingly random and useless until it is organized. When data is processed, organized, structured, or presented in each context to make it useful, it becomes information.”

Davenport, T. H., & Prusak, L. (1998). "Working Knowledge: How Organizations Manage What They Know." Harvard Business School Press.

ETL

“ETL stands for Extract, Transform, Load, a process that collects data from multiple sources, modifies it to meet business requirements, and then stores it in a database or warehouse for analysis. It is a critical component of data management and business intelligence systems.”

Kimball, R., & Caserta, J. (2004). The Data Warehouse ETL Toolkit: Practical Techniques for Extracting, Cleaning, Conforming, and Delivering Data. Wiley.

ETL; Pipeline

“An ETL pipeline is a sequence of processes that automate the extraction of data from source systems, its transformation to align with business rules, and its loading into a data warehouse or other target system.”



Kimball, R., & Caserta, J. (2004). *The Data Warehouse ETL Toolkit: Practical Techniques for Extracting, Cleaning, Conforming, and Delivering Data*. Wiley.

ETL; Task:

“An ETL task represents a specific operation in the extract, transform, and load process, such as retrieving data from a source, applying transformation logic, or writing data to a target system. Tasks are the fundamental components of an ETL pipeline.”

Inmon, W. H., & Linstedt, D. (2014). *Data Architecture: A Primer for the Data Scientist*. Morgan Kaufmann.

API

“An API is a set of programming instructions and standards for accessing a web-based software application or tool. It allows developers to use predefined methods to interact with other software systems and services, facilitating integration and functionality expansion.”

Fielding, R. T., & Taylor, R. N. (2000). Principled Design of the Modern Web Architecture. *ACM Transactions on Internet Technology*, 2(2), 115–150. <https://doi.org/10.1145/514183.514185>

Data; aggregation

“Data aggregation is the process of gathering and summarizing raw data from various sources to create structured and meaningful datasets for analysis and reporting.”

Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly*, 36(4), 1165–1188. <https://doi.org/10.2307/41703503>

Data; Privacy

“Data privacy involves the appropriate handling of personal information to protect individuals' confidentiality, ensure proper use of data, and comply with regulatory and ethical standards.”

Solove, D. J. (2006). A Taxonomy of Privacy. *University of Pennsylvania Law Review*, 154(3), 477–560. <https://doi.org/10.2307/40041279>

Web scraping

“Web scraping refers to the use of automated tools to collect and process data from the web. It involves accessing a website's underlying HTML code, extracting relevant information, and transforming it into a structured format.”

Mitchell, R. (2018). *Web Scraping with Python: Collecting More Data from the Modern Web*. O'Reilly Media.

Python; Jupyter Notebook

“Jupyter Notebook is an open-source tool for interactive computing that combines live code, text, and visualizations into a single document. It is widely used for data analysis, visualization, and educational purposes.”



Kluyver, T., Ragan-Kelley, B., Pérez, F., et al. (2016). Jupyter Notebooks – a publishing format for reproducible computational workflows. In F. Loizides & B. Schmidt (Eds.), *Positioning and Power in Academic Publishing: Players, Agents, and Agendas*. IOS Press. <https://doi.org/10.3233/978-1-61499-649-1-87>

Python; Dataframe

“A DataFrame is a two-dimensional labeled data structure with columns of potentially different types. It is similar to a spreadsheet or SQL table and is one of the most important data structures in the pandas library.”

McKinney, W. (2017). *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython*. O'Reilly Media.

WGS84 standard

“The WGS 84 reference system provides a consistent global framework for geospatial data, defining the Earth's shape, orientation, and coordinate system. It is the standard for GPS operations and ensures interoperability across geospatial technologies.”

Defense Mapping Agency. (1984). *Department of Defense World Geodetic System 1984: Its Definition and Relationships with Local Geodetic Systems*. Technical Report 8350.2. U.S. Department of Defense.

Data License

“A data license provides users with guidelines about how they can access, modify, or share a dataset. For example, the Open Database License (ODbL) allows users to freely use and modify a database, provided they credit the source and share derivative works under the same license.”

Open Data Commons. (2009). *Open Database License (ODbL)*. Retrieved from <https://opendatacommons.org/licenses/odbl/>

Open Database License (ODbL)

“ODbL allows anyone to share, modify, and use a database provided that they attribute the creators, share derivative databases under the same terms, and keep access open to the data.”

Open Data Commons. (2009). *Open Database License (ODbL)*. Retrieved from <https://opendatacommons.org/licenses/odbl/>

F.A.I.R principles

“The F.A.I.R. principles are guidelines for making data Findable, Accessible, Interoperable, and Reusable. These principles ensure that data can be effectively shared and reused in scientific research, enhancing transparency and collaboration.”

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., et al. (2016). “The FAIR Guiding Principles for scientific data management and stewardship.” *Scientific Data*, 3, 160018. <https://doi.org/10.1038/sdata.2016.18>

Map



This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 101016858.

"A map is a graphic representation of the environment. It is a tool designed to convey spatial relationships, whether they be physical, social, economic, or cultural, through a simplified depiction of the features of the Earth or other surfaces."

Harley, J. B., & Woodward, D. (1987). *The History of Cartography, Volume 1: Cartography in Prehistoric, Ancient, and Medieval Europe and the Mediterranean*. University of Chicago Press.

Mercator projection

"The Mercator projection is a cylindrical map projection designed for navigation that preserves local angles and shapes but distorts the size of landmasses as latitude increases from the equator to the poles."

Snyder, J. P. (1987). *Map Projections—A Working Manual*. U.S. Geological Survey Professional Paper 1395.

Machinery

"Machinery refers to devices or systems composed of interrelated parts that work together to perform a specific task, particularly in industrial or manufacturing processes. They may be manually operated, semi-automated, or fully automated, and are designed to transform raw materials into usable components or products."

Groover, M. P. (2015). *Fundamentals of Modern Manufacturing: Materials, Processes, and Systems*. Wiley.

Digital machinery

"Digital machinery refers to equipment that combines traditional mechanical systems with digital components, such as sensors, processors, and communication technologies, to enable automation, monitoring, and enhanced operational efficiency. It forms the backbone of smart manufacturing and modern industrial processes."

Kagermann, H., Wahlster, W., & Helbig, J. (2013). *Recommendations for Implementing the Strategic Initiative INDUSTRIE 4.0. Final Report of the Industrie 4.0 Working Group*. Acatech.

FabLab

"A FabLab is a technical prototyping platform for innovation and invention, providing access to tools for digital fabrication. It is simultaneously a place for learning and innovation: a place to play, to create, to learn, to mentor, to invent. FabLabs are closely tied to the global FabLab network, which allows for collaboration and knowledge sharing."

Gershenfeld, N. (2005). *Fab: The Coming Revolution on Your Desktop—From Personal Computers to Personal Fabrication*. Basic Books.

DIY, Do-It-Yourself

"DIY refers to activities in which individuals create or modify products and structures on their own, typically without professional expertise, fostering a culture of creativity and self-reliance."



Kuznetsov, S., & Paulos, E. (2010). Rise of the Expert Amateur: DIY Projects, Communities, and Cultures. *Proceedings of the 6th Nordic Conference on Human-Computer Interaction*, 295–304. <https://doi.org/10.1145/1868914.1868950>

Maker

“Makers are individuals who create and innovate, combining traditional crafting methods with modern technologies, often within a community that emphasizes open collaboration and learning.”

Dougherty, D. (2012). The Maker Movement. *Innovations: Technology, Governance, Globalization*, 7(3), 11–14. https://doi.org/10.1162/INOV_a_00135

Makerspace

“A makerspace is a collaborative workspace for making, learning, exploring, and sharing that uses high-tech to no-tech tools. These spaces are open to everyone, providing a place for hands-on experimentation and creative problem-solving.”

Sheridan, K., Halverson, E. R., Litts, B., Brahms, L., Jacobs-Priebe, L., & Owens, T. (2014). Learning in the Making: A Comparative Case Study of Three Makerspaces. *Harvard Educational Review*, 84(4), 505–531. <https://doi.org/10.17763/haer.84.4.brr34733723j648u>

KoboToolBox

“KoBoToolbox is a free, open-source tool for mobile data collection that allows users to create and manage surveys, even in offline settings. It is widely used in humanitarian and development projects for data collection, monitoring, and evaluation.”

Harvard Humanitarian Initiative. (2023). KoBoToolbox Documentation. Retrieved from <https://support.kobotoolbox.org/>



Introduction

The Map of Machinery project is a tool designed to empower innovation by mapping and connecting small-scale manufacturing capabilities. By finding reliable data sources, performing surveys, designing maps, researching use cases, and analysing data this initiative fosters collaboration and participation in the emerging paradigm of Distributed Manufacturing—a system that reimagines how goods are designed, produced, and shared on a global scale. This section of the document explores the reasoning, scope, and key components behind this impactful project. We begin by exploring the project into its components and definitions:

What is a Map?

"A map is a graphic representation of the environment. It is a tool designed to convey spatial relationships, whether they be physical, social, economic, or cultural, through a simplified depiction of the features of the Earth or other surfaces."¹ In the context of the Map of Machinery, we employ Mercator projection² maps, as it's a common, familiar and technically fast representation of the world, compared to 3D representations. We also make use of the following graphic elements for the representation of locations:

- **Points:** To mark precise locations, such as individual makerspaces.
- **Lines:** To represent connections between locations.
- **Areas:** To outline regions or boundaries.

Henceforth, the map of machinery will serve for the correlation of geographical distribution and manufacturing capabilities and machinery.

What is Machinery?

"Machinery refers to devices or systems composed of interrelated parts that work together to perform a specific task, particularly in industrial or manufacturing processes. They may be manually operated, semi-automated, or fully automated, and are designed to transform raw materials into usable components or products."³

For the map of machinery, we classified machines as:

- **Traditional:** Operated manually or semi-manually, often relying on human oversight.
- **Digital:** Automated or semi-automated tools controlled via software or digital interfaces.

¹ Harley, J. B., & Woodward, D. (1987). *The History of Cartography, Volume 1: Cartography in Prehistoric, Ancient, and Medieval Europe and the Mediterranean*. University of Chicago Press.

² Snyder, J. P. (1987). *Map Projections—A Working Manual*. U.S. Geological Survey Professional Paper 1395.

³ Groover, M. P. (2015). *Fundamentals of Modern Manufacturing: Materials, Processes, and Systems*. Wiley.



Categorizing also serves the purpose of mapping a wider social-economical range of manufacturing workshop spaces and processes. Digital manufacturing machinery is non-easily accessible as it is tied to the availability of training, energy infrastructure and economic resources. The maker scene has a broader community after the COVID-19 global response needed the manufacturing of facemask and protective equipment to slow down the spread of the virus. The incorporation of sewing machines on the map in regions of Uganda and Kenya highlighted the need to account on textile related machines. In the following sections, we will explore this scenario of collecting data of SMEs and its validation process. On the other side, digital manufacturing machines are tied to DIY, Small Manufacturing, repair and prototyping services, at innovation hubs, universities, makerspaces, and online manufacturing platforms.

What is DIY, repair communities, Small Manufacturing and prototyping spaces?

DIY stands for Do-It-Yourself; it's the foundational motto of the Maker movements, like the Free and Open-Source Software, the Open-Source Hardware, and skills learning initiatives, where cooperation and exchange have opened the possibilities of developing solutions that won't fit commercially available products.

A repair community comes with the need of refactoring, fixing and replacing parts of equipment that has long served its designated life cycle or has become too expensive or impossible to source due to supply chain related complications, like the ones experienced during the COVID-19 pandemic.

Small Manufacturing Businesses are an integral part of the economy in developing countries, often are equipped by traditional machinery and perform custom manufacturing of goods, tailoring and small batch manufacturing. Their main purpose is to supply for the local market and most of their designs and raw materials are dependable of the cultural, economic and geographical contexts.

Finally, citing "A makerspace is a collaborative workspace for making, learning, exploring, and sharing that uses high-tech to no-tech tools. These spaces are open to everyone, providing a place for hands-on experimentation and creative problem-solving."⁴

The focus of the initial data assessment was on FabLabs, a type of makerspace with a global community, with defined machine inventory and a platform that shares the data openly. All these spaces unite and form groups or communities, for knowledge exchange, financial cooperation, support and differentiation.

What are Distributed Manufacturing Networks?

Distributed Manufacturing Networks are at the heart of this project. These networks consist of individuals or organizations engaged in small-scale manufacturing or prototyping, often united by a shared mission, geography, or purpose.

⁴ Sheridan, K., Halverson, E. R., Litts, B., Brahms, L., Jacobs-Priebe, L., & Owens, T. (2014). Learning in the Making: A Comparative Case Study of Three Makerspaces. Harvard Educational Review, 84(4), 505–531.
<https://doi.org/10.17763/haer.84.4.brr34733723j648u>



Our data search revealed a rich diversity of regionalized databases, each with unique focuses, including:

- Handcrafting and artisan networks.
- COVID-19 initiatives like facemask production.
- Innovation hubs with prototyping facilities.
- Maker-space networks.
- Public libraries with maker capabilities.
- Universities with FabLab.

These networks represent the decentralized ethos of Distributed Manufacturing: local production with global impact, facilitated by collaboration and shared resources. A more in-depth review of each data source is detailed in the following section of data assessment.

1. The map of machinery

1.1. What is the Map of Machinery?

The Map of Machinery is a Geographic Information System (GIS)⁵ developed to locate, analyse, and enhance the accessibility of small-scale manufacturing capabilities. As part of a broader Distributed Manufacturing Network, it serves as a bridge between makers, communities, and stakeholders across Africa and Europe.

The project development involved the indexing of similar resources, the location of reliable data in them, the search for use cases for the map, the design of a data model that accounts for privacy regulations and transparency, a system for the data processing, and the deployment of interactive maps.

1.2. What are the use cases proposed for the map of machinery?

1.2.1. Finding Manufacturing Sites for Open-Source Projects

Open-source hardware designers can locate suitable facilities to manufacture their designs, ensuring access to the right equipment, expertise, and materials. This enables efficient scaling of innovation and supports localized production for global challenges.

1.2.2. Employment Creation Guidelines

Communities can leverage distributed manufacturing to stimulate economic growth and create jobs. By providing data-driven insights, the Map of Machinery helps establish safe, inclusive manufacturing nodes tailored to local capabilities and needs.

⁵ Goodchild, M. F. (1992). Geographical Information Science. *International Journal of Geographical Information Systems*, 6(1), 31–45. <https://doi.org/10.1080/O2693799208901893>



1.2.3. Maker-Space Learning and Setup Guidelines

Aspiring makers and organizations gain access to best practices and resources for setting up and managing makerspaces. These guidelines cover essential aspects such as equipment selection, safety protocols, and community engagement strategies.

1.2.4. Rapid Response Procurement Tool

In emergencies, the platform enables organizations to quickly identify and procure critical supplies, ensuring smooth coordination and efficient response during crises.

2. Data assessment

The first data trial of the project was exploring existing solutions to understand the landscape of maps for distributed manufacturing. For example, the FabLab Network Map (<https://fablabs.io/labs>) showcases FabLab locations worldwide, demonstrating the value of maps in finding and understanding the reach of their organisation. While the data extraction process from this platform posed challenges, extensive API documentation revealed a useful endpoint for gathering data (<https://api.fablabs.io/O/labs.json>), it directs to a JSON file that holds the map markers information as text. JSON is often used as a data exchange format.

Data came in different formats, some intended for its transport others were limited offering a machine-readable output. A task of data scraping was conducted to extract it, like in the case of MakerTour – an NGO based in France – that collected and mapped its data and published it as a Mercator Map on its website.

The intention of collecting data was to merge and consolidate all data sources in one source, to find points with multiple references, and the completion of data from as many different sources as possible. Some organisations have a specific interest and only the data related to it is collected and mapped. The most extensive data source found came from FieldReady, an NGO covering the manufacturing of facemasks during the COVID-19 pandemic procurement of resources for humanitarian purposes. They covered their data following the Open-Know-Where metadata specification, an open-source contribution by the Internet of Production Alliance. The simplest data source found is a directory of libraries across the United States of America, with accounts to Name of the place, A short description, a location Marker and an URL for more information. This map was shared using Google Maps.

Some organisations will have their own resources to deploy a self-hosted map in their webpages, others will use services like Google Maps, LeafletJS and others will have their own API. This also highlights the availability of IT resources by some organisations and the need to find a cost-effective, maintainable and performative solution for the Map visualisation.



Building on these foundations, the Map of Machinery aggregates data from diverse sources, creating a centralized platform that supports the mAKE consortium of partners and the broader maker movement in Africa and Europe.

On second and third trials, we conducted data collection by surveying using tools like KOBOToolbox, a mobile-offline form creator and data collection system. A set of tools were facilitated, and a sponsored data collection process was conducted with an open call for applications ([OKW Data Awards](#)) and a selection process.

[Open Know-Where \(OKW\)](#) was launched in April 2021 as a metadata specification to document and share information about the location of manufacturing facilities and capabilities globally.

The **OKW Data Awards** provides financial support for projects that integrate existing datasets or collect new data related to the mapping of manufacturing capabilities. Round 1 of the OKW Data Awards provided the experience of applying the standard in community projects, bringing closer the goal of OKW for being an initiative to help map manufacturing facilities and machinery for use in decentralized manufacturing. The objective for 2023 was to map at least 150,000 points of interest (i.e.: facilities and machinery), under the OKW initiative. This mapping, in addition to benefiting community makers in finding where to produce items, contributed to the development of the infrastructure for the distributed production of hardware: a global map of facilities, machines, and tools, based on shared open data and open APIs. The recipients of this award have demonstrated exceptional skills and expertise in applying OKW to the mapping of manufacturing facilities and machinery in support of the global community.

Takeouts of the data assessments:

- Many organizations depend on third-party tools for mapping and data hosting, rather than managing these resources independently.
- Data collection methods are often tailored to the organization's specific focus or interests, leading to diverse but inconsistent models.
- A significant number of organizations lack a formal data quality assurance process, impacting the reliability and accuracy of their datasets.
- Data sharing and exchange are limited, with few organizations adopting standardized formats or protocols for open data exchange.
- Multilingual accessibility is largely absent, restricting the usability of data for diverse audiences.

The reliability of collected data varies significantly across organizations due to differences in their validation and collection processes:

- Some organizations ensure high reliability by deploying members to travel abroad and conduct in-person interviews with those involved in the manufacturing process. This approach provides valuable context and insights into capabilities but becomes less cost-effective as the geographic area of data collection expands.



- Other organizations implement peer validation processes, requiring confirmation from at least two active validator members. While this method ensures a degree of accuracy, its effectiveness depends on the validators' expertise and adherence to guidelines.
- Certain organizations rely on tools that automatically capture extensive metadata about the location and activity of the data collector. While these tools add an additional layer of traceability, they may not provide the same depth of contextual understanding as in-person interviews.

The process of identifying open, publicly available data sources within this project scope was conducted following the keywords and topics during web search as: “makerspaces, fablabs, hackerspaces, DIY workshops, prototyping spaces, 3d printing–laser cutting machining services, and innovation–hubs.” Over a span of 8 months, different sites, documents and files were reviewed, following the criteria as follows: “up-to-date, relevancy, accuracy, formatting, endorsements, and originality”. The data covered world-wide regions, with a priority to data in Africa and Europe. The following list is a capture for an ongoing data collection process.



Table 1. Data source tables integrated on the Map of Machinery ETL system:

Organization	URL	License	Records	ID
The Fab Foundation, FabLab Network and partners	https://www.fablabs.io/labs	Limited	2581	2
Bay Area Makerspaces Upbeatlabs	https://www.upbeatlabs.com/	Open (Unspecified)	58	3
University of Michigan 3DLab	https://library-maker-culture.weebly.com/	Open (Unspecified)	116	4
MakerTour	https://www.makertour.fr/map	CC BY-SA 4.0	156	5
Make Works, FabLab Barcelona and partners	https://make.works/companies	CC BY-NC 4.0	359	6
Offene Werkstaetten, and partners	https://www.offene-werkstaetten.de	Limited	512	7
Bay Area Makerspaces UC Berkeley	https://jacobsinstitute.berkeley.edu/other-spaces/	Open (Unspecified)	36	8
Hackerspaces.org	https://hackerspaces.org	Open (Unspecified)	2518	9
Makery	https://www.makery.info/map-labs/	CC BY-NC-SA 3.0	533	10
Makezine, Makerspace.com	https://makerspace.com	CC BY-NC-SA 4.0	782	11
Field Ready Inc. NeedsList and partners	https://www.fieldready.org	Limited	2185	12
IoP Data Awards R1 & R2	WIP	ODbL	400	13
IoP IMA	WIP	ODbL	9	14
sphere.diybio.org	https://sphere.diybio.org	ODbL	63	15
		Total	*10,308	

An overall count of *20,000 records was aggregated, and after filtering out were obtained *10,308 unique records with a minimal set of attributes “name, (latitude, longitude, or address), (capability or machine), and (URL, e-mail or phone number)”. The discarded data didn’t complete the filtering criteria. In the following sections the data architecture will specify how data privacy is ensured by design in the data aggregation process.

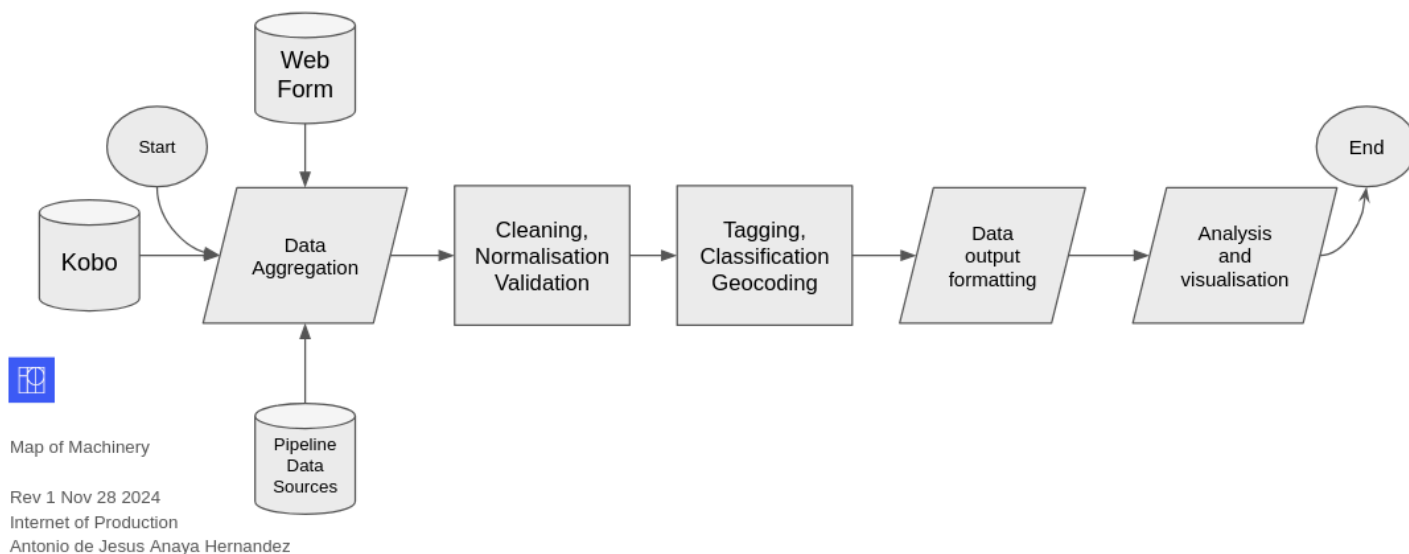


The process of data filtering followed the criteria of "location completeness, contact information, open or active, valid name, located in land, non-duplicated, and does not contain; null, none, empty, unformatted and noise data." Data filtering was necessary from individual sources to keep a curated dataset for increased reliability. A clustering process was part of the aggregation strategy to identify records with references across data sources. In a way that all referenced data is concatenated.

The process of extracting data was programmed following the criteria: "reproducible, automated, without-side effects, maintainable, and low-cost". For that a series of computer programs were written in Python language version 3.12, for the data extraction, the proofs of concept are demonstrated in Jupyter Notebooks published on the publicly available version control repository "okw_data_management" at https://github.com/iop-alliance/okw_data_management".

The data source formats found were: "JSON, HTML, JavaScript LeafletJS, KML, CSV, SQL". Each format has a method of extraction, some extraction processes required the filtering and parsing of HTML content in a process known as Web-Scraping. The overall process is defined as an ETL (Extract, Transform, Load) system, visualised as:

DevOps ETL Infrastructure for the Map of Machinery



This process was designed during the early stages of the Map of machinery project for the aggregation of data of manufacturing capabilities. In the following section the deployment of this ETL system with Metaflow is detailed.



2.1. Details of data sources

FabLab Network, Fab Foundation, and partners

The screenshot shows the FabLabs.io website interface. At the top, there is a search bar with the text "Search, anything" and a navigation menu with links for Labs, Orgs, Machines, Jobs, Fab Academy, Events, Network, Projects, and Forum. There are also "Sign in" and "Sign up" buttons. Below the navigation, there are "List view" and "Map view" buttons. A filter section includes "Filter by Country", "Filter by activity status", "Filter by lab tags", and a "Show: 25" dropdown with a "Filter" button. A "+ Add Lab" button is also present. The main content area shows a list of Fab Labs on the left and a map on the right. The list includes:

- Fab Lab Afghanistan (Jalalabad, Nangarhar)
- Fab Lab Manchester (Manchester, Greater Manchester, United Kingdom)
- FabLab Valencia (Valencia, Valencian Community, Spain)
- HappyLab Wien (Vienna, Vienna)
- fablab IMAL

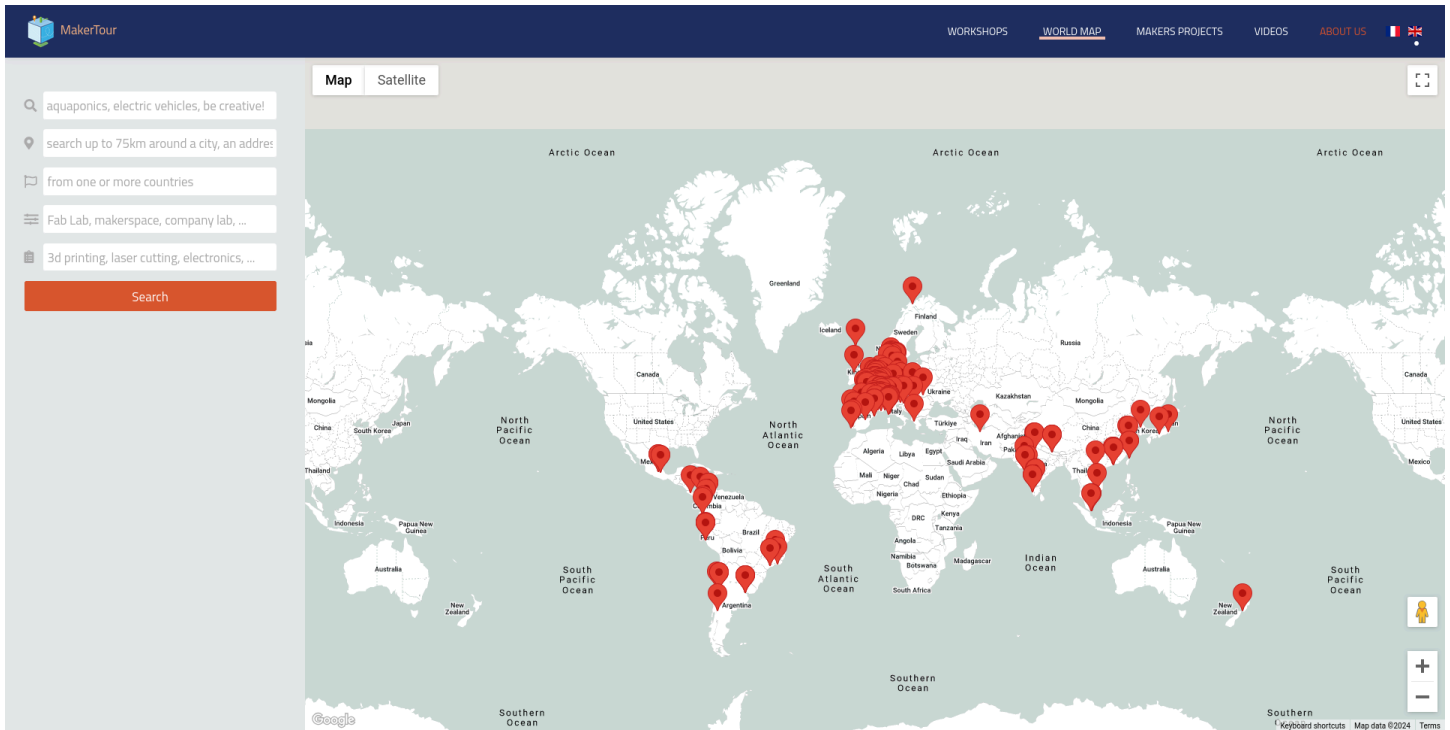
The map shows a global view with a FabLab icon placed in Afghanistan. A pagination bar at the top of the list shows "1 2 3 4 5 ... Next Last". A cookie notice at the bottom states: "This website uses cookies to ensure you get the best experience on our website. More info" with a "Got it!" button.

FabLabs.io is the online social network of the international FabLab community. More than 1750 FabLabs exist all over the world. The FabLab network connects global communities of learners, educators, technologists, researchers, makers and innovators, spread in 100 countries and 24 time zones. All FabLabs share common tools and processes building a global network and creating a distributed laboratory for research and invention.

Data columns: ['id', 'name', 'kind_name', 'parent_id', 'blurb', 'description', 'slug', 'avatar_url', 'header_url', 'address_1', 'address_2', 'city', 'county', 'postal_code', 'country_code', 'latitude', 'longitude', 'address_notes', 'phone', 'email', 'capabilities', 'activity_status', 'links']



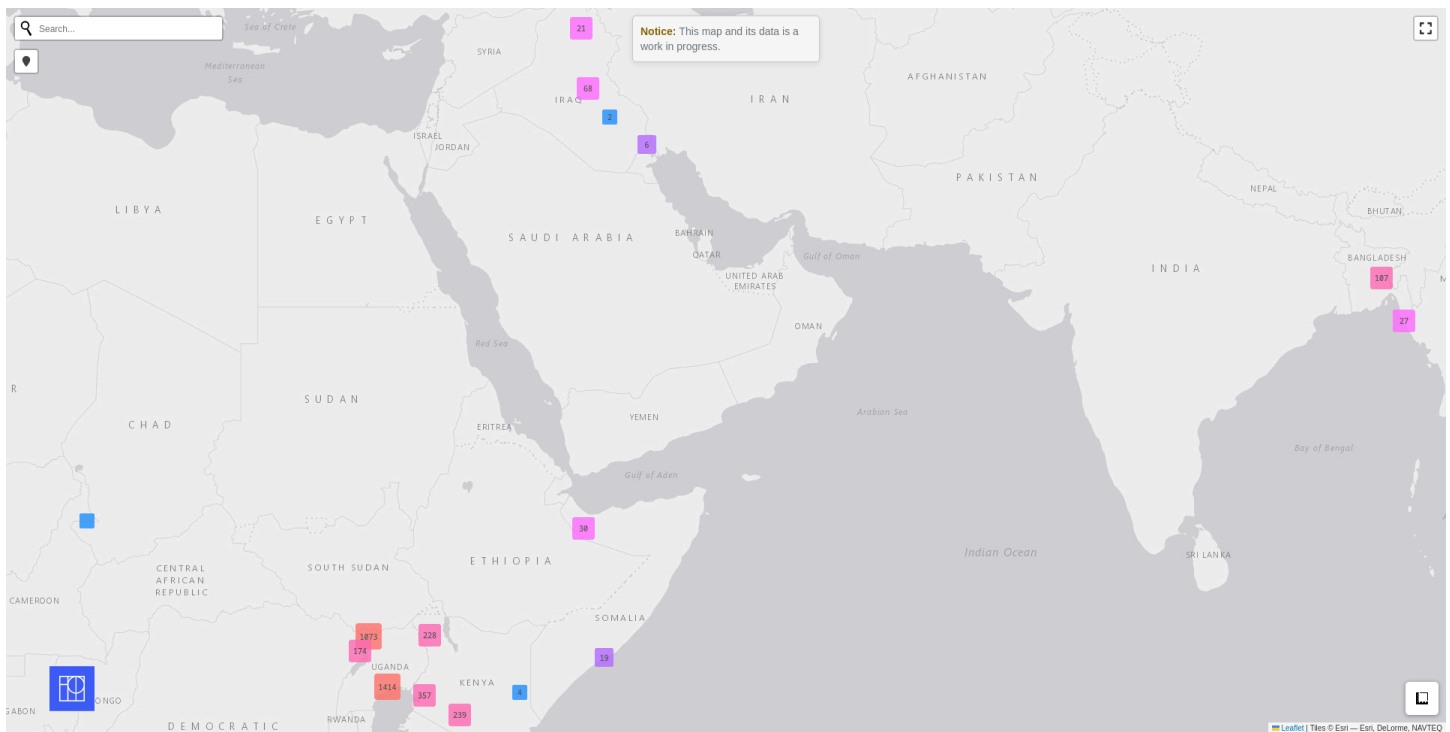
Maker Tour, and partners



Makertour is a French non-profit organisation. Since 2015, MakerTour has launched expeditions to explore workshops and meet local communities worldwide. The projects and the places they discover are shared, documented on <https://makertour.fr/projects>

Data columns: ['name', 'city', 'country', 'tags', 'record_source_url', 'latitude', 'longitude']



Field Ready Inc, NeedsList, and partners

Field Ready is a non-governmental, non-profit organisation aiming to meet humanitarian needs by transforming logistics through technology, innovative design and engaging people in new ways.

Field Ready, a multinational group of non-governmental organisations, employs the latest best practices and advanced thinking from the fields of relief and development, human-centred design, engineering and various forms of making and manufacturing. They are mostly present in Nepal, Syria, South Pacific, Iraq and in the United States of America, bringing local manufacturing to challenging places, focusing on reconstruction to help the communities.

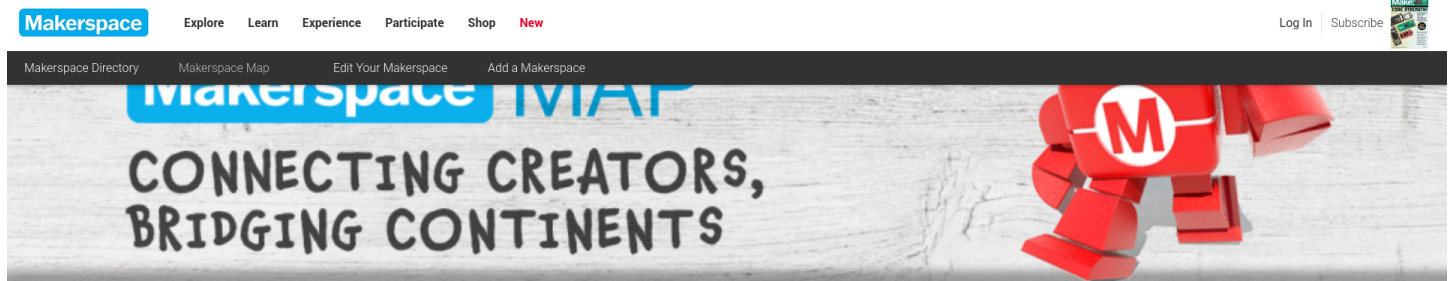
Data columns: ['fid', 'start', 'end', 'today', 'enumerator', 'name', 'country', 'addr_district", 'dbdistrict', 'governorate', 'addr_village', 'address', 'addr_street', 'number', 'postcode', 'location-Latitude', 'location-Longitude', 'location-Altitude', 'location-Accuracy', 'Type', 'owner', 'Bio', 'contact', 'email', 'social', 'affiliation', 'partner_funder', 'certifications', 'facility_status', 'Working_hours', 'working_days', 'date_founded', 'storage_capacity', 'size_floor_size', 'loading_dock', 'back_up_generator', 'uninterrupted_power_supply", 'wheelchair_access', 'road_access', 'headcount', 'maker', 'product', 'batch_size', 'equipment_available', 'model', 'serial_number', 'Wikipedia_URL,, 'Manufacturing_process', 'Wikipedia_URL_of_the_anufacturing_process', 'Materials_used', 'Plastic_s_type', 'Metal_s_type', 'Wood_type', 'Elastomer_s_type', 'Ceramics_type', 'Electronics_type', 'Others', 'maintenance_schedule']

1994
Africa
5

149
Asia
2



Makezine, makerspace.com, and partners



Makerspaces, also known as hackerspaces or fab labs are more than just workshops—they are hubs of inspiration, experimentation, and learning.

Maker Faire Community Round

INVEST IN THE NEXT GENERATION; TOGETHER, LET'S SHAPE THE FUTURE.

[LEARN MORE](#)

Makerspace Directory, is a platform developed by the Make: Makezine, a comprehensive guide to the growing community of makerspaces around the globe. This directory is a gateway to discovering innovative communities and collaborative environments. In this directory, you'll find a community curated collection of makerspaces spanning continents, cultures, and disciplines. Check back often as more spaces and features are being added.

Data columns: ['id', 'link', 'name', 'image', 'logo', 'type', 'address', 'city', 'state', 'country', 'zip', 'lat', 'long', 'desc', 'website', 'email', 'social']



Hackerspaces.org, and partners

List of Hacker Spaces

This is a comprehensive, user-maintained list of all active hackerspaces throughout the world.

There are currently 2518 hackerspaces listed in this wiki, 812 of them are marked as active and 352 as planned. We have also a [list of planned Hacker Spaces](#), as well as a [list of All hackerspaces](#) around the globe - including those still in building process - or already closed. Realtime data of some active hackerspaces is available at the [SpaceAPI](#) project.

This page is cached due to its long rendering time. If you want to flush the cache, [click here](#).

hackerspace

hackerspace	Country	State	City	Website	Last Updated
Cascadia Makerspace	United States of America	Oregon	Hillsboro	https://cascadiamakerspace/	20 November 2024
Left Click	Slovenia	Ljubljana	Ljubljana	https://leftclick.si	18 November 2024
Hackerspace Lodz	Poland	Łódzkie	Łódź	https://hs-lodz.pl	17 November 2024
Glasgow Hackerspace	United Kingdom	Scotland	Glasgow		14 November 2024
KanGouly.org	Tunisia	Tunisia	Tunis	https://kangouly.org	12 November 2024
Cascadia Makerspace	United States of America	Colorado	Windsor	http://cascadiamakerspace.com/	12 November 2024

Hackerspaces (also hacklabs, hackerspaces) are community-operated physical places, where people share their interest in tinkering with technology, meet and work on their projects, and learn from each other.

Hackerspaces.org is an informal volunteer network of such spaces, maintaining community services – including an open wiki for everyone who wants to share their hackerspace information, stories and questions. From around the world, hackers meet on the libera.chat IRC channel #hackerspaces that is bridged to #chat:hackerspaces.org on [matrix].

Data columns: ['name', 'latitude', 'longitude', 'url']



Verbund Offener-Werkstätten, and partners

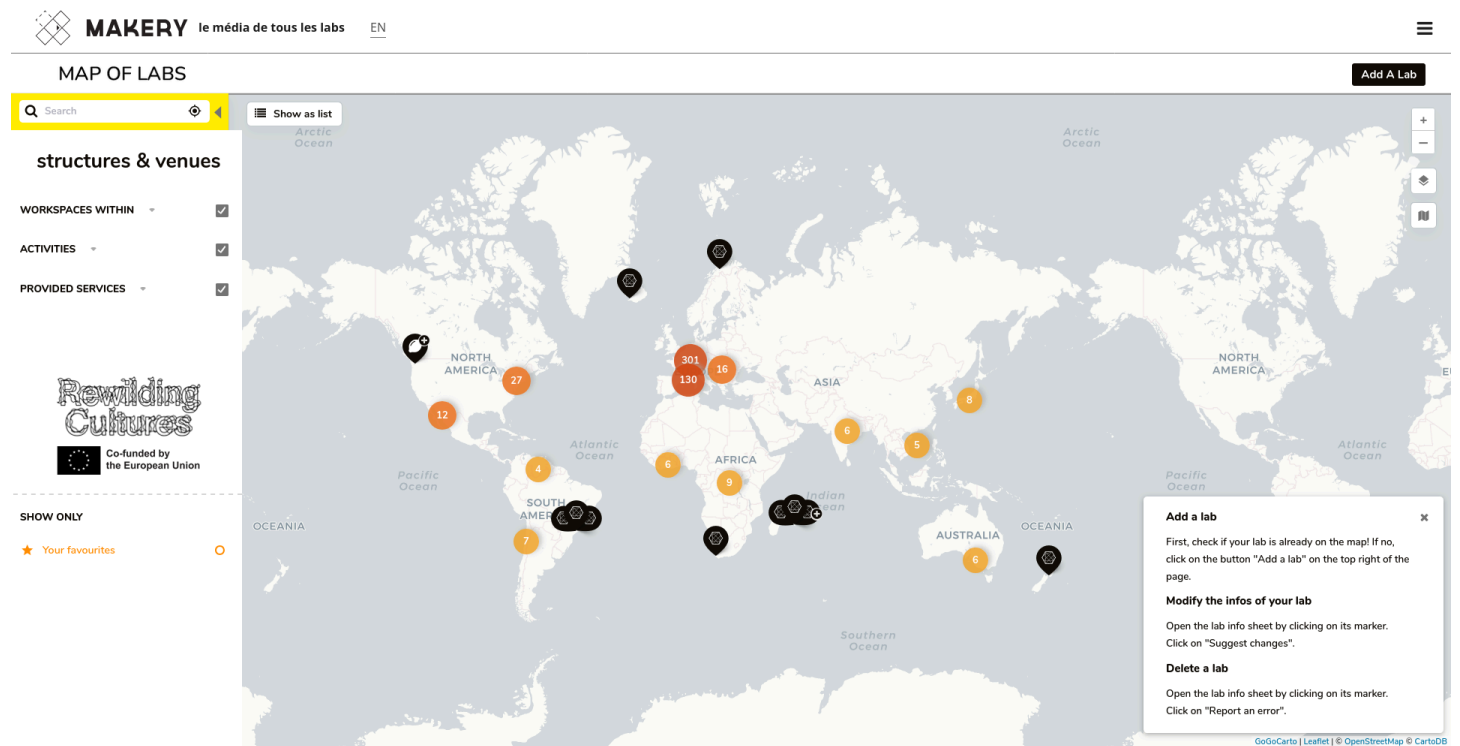
Verbund Offener Werkstätten (VOW) is the association of the German Open Workshops. Open workshops are places of craftsmanship, computer-controlled manufacturing processes and digital technologies. They are spaces to share knowledge, tools and machines, technology and materials.

Data columns: ['name', 'img', 'uid', 'url', 'lat', 'lng', 'street', 'zip', 'city', 'web', 'country', 'cats', 'aai', 'icm', 'street_nr']

513
Europe
6



Makery, and partners



Makery.info is an online magazine, bilingual French/English, free and ad-free, covering the Do-It-Yourself and creative communities scene since 2014. Makery carries the voice of cultural players, intellectuals, artists and more generally all those who are part of social innovation, think and act for a more ecological, inclusive and open world.

Directed by Anne-Cécile Worms (art2m.eu), Makery is also a partner in several international cultural programs and networks, and involved in several publications, residency programs, research-action programs, artistic and cultural events.

Data columns: ['id', 'm_id', 'region_ids', 'name', 'address', 'background', 'description', 'intro', 'contact_name', 'contact_email', 'contact_phone', 'contact_jobtitle', 'flickr', 'facebook', 'instagram', 'large_run', 'lat', 'lng', 'linkedin', 'location', 'medium_run', 'minimum_order', 'number_of_staff', 'file_types', 'image_bucket', 'photo1', 'photo2', 'photo3', 'photo4', 'photo5', 'photo6', 'photo7', 'photo8', 'photo9', 'pinterest', 'portrait', 'sample_production', 'short_run', 'soft_delete', 'title', 'top_image', 'slug', 'turnaround_time', 'twitter', 'video_link', 'website', 'youtube', 'year_founded', 'film_ready', 'created_at', 'updated_at', 'url']

207
Europe
4

58
Asia
1



Make Works, and partners

Make Works

Manufacturers

Stories

About

Log in

Join free

Explore all manufacturers

Search for manufacturers, materials, processes...



Regions

Industries

Materials

Machines

Processes

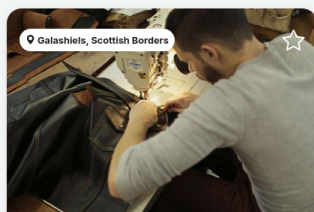
Products

Character tags

Search

[Clear all](#)

347 results



Galashiels, Scottish Borders

Aero Leather Clothing

Aero Leathers design and manufacture high grade clothing, specialising in vin...



Isle of Elgg, Highland

AnNead Hand Knitwear

Jenny Robertson hand knits fine cobweb lace from hand-spun yarn. She makes be...



Glasgow

Andrew Muirhead & Son

Andrew Muirhead are a tannery, supplying high performance low carbon leather ...

Make Works is a global library for local manufacturing, making manufacturing accessible. In the library, a series of open resources are available for people to find local manufacturers, tools and materials. It is free and accessible; it connects people who want to make and repair things. Make Works partners with individuals and organisations across the UK and beyond who want to run their own open-access directory of manufacturing for their local area. The project started with the mapping of manufacturing in Glasgow. In 2019, Make Works was transferred to the Institute of Advanced Architecture of Catalonia (IAAC), Fab City Research Lab Barcelona, and sits within the Fab City's Distributed Design Platform and the Fab City's Global Initiative. Make Works operates in 18 regions.

Data columns: ['id', 'm_id', 'region_ids', 'name', 'address', 'background', 'description', 'intro', 'contact_name', 'contact_email', 'contact_phone', 'contact_jobtitle', 'flickr', 'facebook', 'instagram', 'large_run', 'lat', 'lng', 'linkedin', 'location', 'medium_run', 'minimum_order', 'number_of_staff', 'file_types', 'image_bucket', 'photo1', 'photo2', 'photo3', 'photo4', 'photo5', 'photo6', 'photo7', 'photo8', 'photo9', 'pinterest', 'portrait', 'sample_production', 'short_run', 'soft_delete', 'title', 'top_image', 'slug', 'turnaround_time', 'twitter', 'video_link', 'website', 'youtube', 'year_founded', 'film_ready', 'created_at', 'updated_at', 'url']

207
Europe
458
Asia
1

3. Proof of concept for data collection

A GIT repository containing Jupyter notebooks for the collection and analysis of machinery and facilities is available at:

https://github.com/iop-alliance/okw_data_management/tree/main/notebooks/data_collectors .

These notebooks cover various aspects of data management, including data acquisition processes, machinery cataloguing, and structured data handling. The following is a list of methods identified for data collection from existing GIS systems and databases:

- **Data Extraction via HTTP Requests:**
Extract data from an endpoint that serves a JSON file.
- **Web Scraping:**
Collect data from web pages that lack an API or relevant endpoints.
The HTML content of the web page is downloaded and parsed to extract specific information rendered in the web content.
- **Data Collection from GIS Files:**
Handle data in GIS formats, such as KML (Keyhole Markup Language), Google-Maps, to retrieve geospatial information.
- **Database Queries:**
Retrieve tabular data from relational databases using structured queries.
- **Parsing Online Mapped Data:**
Parse and process geospatial data rendered on online mapping platforms for analysis and integration. Like i.e. extracting JSON data embedded in Leaflet.JS maps.
- **Machinery identification:**
Once the source data is collected, functions are run to identify machinery, following the classification “3d printer, CNC machine, laser cutter and others.”



3.1. Catalogue of machinery

A data analysis of machinery identified within the FabLab Network was performed. The catalogue at <https://fablabs.io/machines> is a dataset of self-submission of data accounting machinery available at FabLabs, with an undefined data model, whereas some content is tagged but most doesn't have a uniform tagging for the classification, as some machines do not have any tagging to differentiate them from 3D printers, CNC millers, laser cutters, etc. A data completion using Natural Language Processing was executed.

The detailed data analysis can be found at:

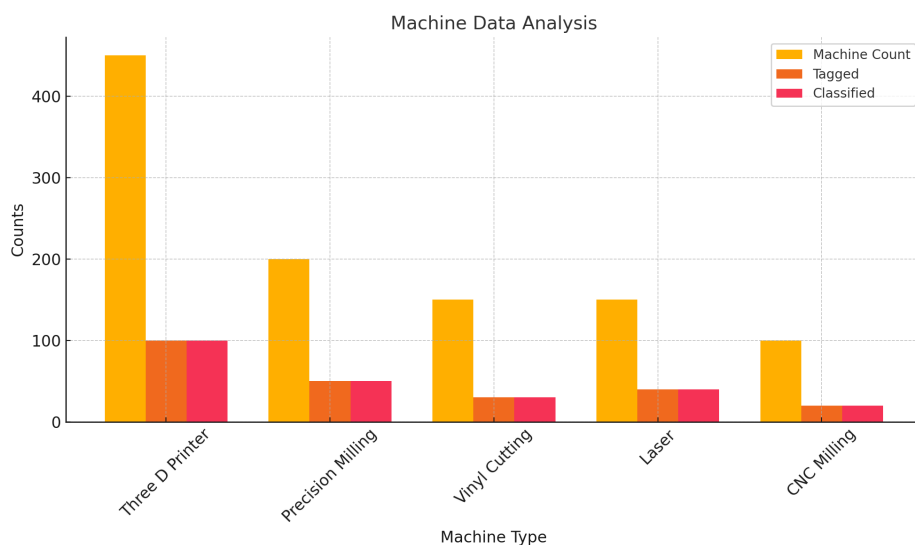
https://github.com/iop-alliance/data_reports/blob/main/notebooks/machines_in_fablabs_world_tags.ipynb

The darker lines represent missing tagging data, for machine classification.



Graph representing the difference between tagged and unclassified machinery data in the FabLabs.io platform: Example of tags: "3d printer", "laser cutter", "CNC machine".

The objective of tagging/labelling data entries is to keep a uniform classification of data for analysis and insights.



A product of this data analysis across sources is the Catalogue of machinery repository, it holds technical specifications of machinery, like sizes, materials and types of manufacturing process for digital machinery.

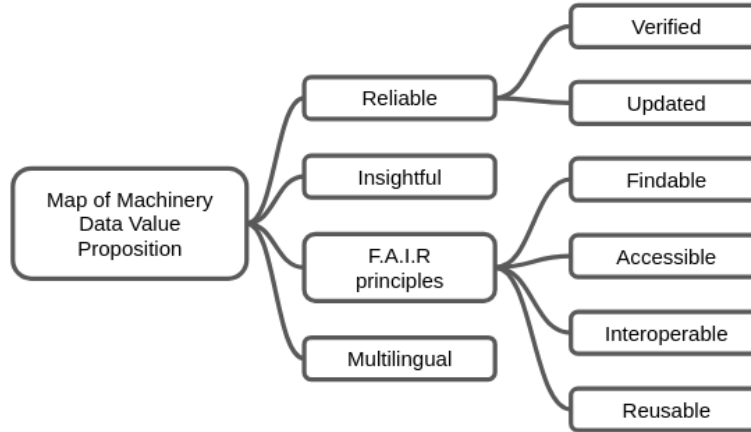
More details at: https://github.com/iop-alliance/okw_machines_catalog



4. Data specification

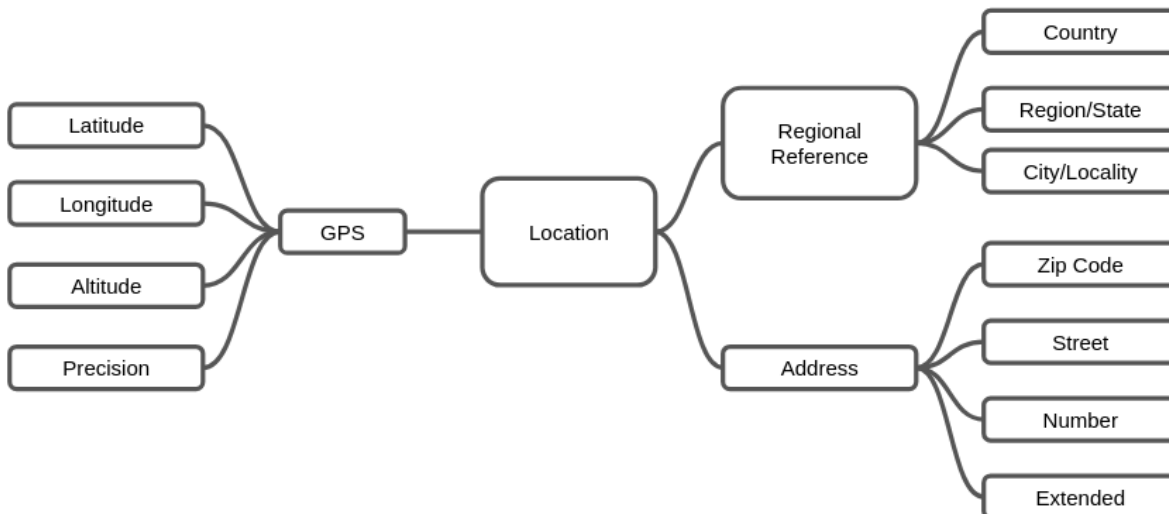
4.1. Data architecture

Designing a data source specification that enables people to engage in distributed manufacturing. Following FAIR principles and contributing to the Open-Source Ecosystem, having the mission to offer useful and localisable data. This initial modelling covers the missing information observed during the data assessment exercise.



To uphold GDPR standards, the data model is built with strong safeguards for protecting personal information. A structured classification system was established to evaluate privacy sensitivity, ensuring that data collection remains purposeful and compliant while minimizing risks to individuals:

- **Highly Safe:** Regional references and web handlers (e.g., URLs) are considered the least sensitive types of data and can be openly shared with minimal privacy concerns.
- **Moderately Sensitive:** GPS coordinates and physical addresses offer more specific locational data and require careful handling to prevent potential misuse.
- **Highly Sensitive:** Emails and phone numbers are classified as critical data points, requiring the highest level of protection to ensure confidentiality and prevent unauthorized access.



This approach is grounded in GDPR principles, particularly the emphasis on data minimization and user consent. It ensures that only the essential data needed to achieve the project's objectives is collected, processed, and stored, while respecting the rights and privacy of individuals.

The data found during the data assessment and collection has provided the insight of aggregating data of elements interacting during a distributed manufacturing network production process. Like: Facilities, Equipment/Machinery, Services and Skills. This work of the Distributed Manufacturing Resources is not covered entirely in this report, we will limit the scope to explain the Map of Machinery data.

For more details refer to the data specification repository:

https://github.com/iop-alliance/okw_data_management/blob/main/docs/okw_v1_schema_specification.md



OKW data schema diagram.

C FindableProject
<ul style="list-style-type: none"> o project_id: int [unique, primary key] o resource_id: int [ref: > DistributedResource.resource_id] o name: string o description: string o source_id: int [ref: > DatabaseSourceMetadata.source_id]

C Address
<ul style="list-style-type: none"> o address_id: int [unique, primary key] o street: string o number: string o locality: string o region: string o postcode: integer o country: string o latitude: float o longitude: float

C ManufacturingSkill
<ul style="list-style-type: none"> o skill_id: int [unique, primary key] o resource_id: int [ref: > DistributedResource.resource_id] o name: string o bio: string o skills: string o source_id: int [ref: > DatabaseSourceMetadata.source_id]

C Contact
<ul style="list-style-type: none"> o contact_id: int [unique, primary key] o social: string o email: string o website: string

C ManufacturingService
<ul style="list-style-type: none"> o service_id: int [unique, primary key] o resource_id: int [ref: > DistributedResource.resource_id] o type: string o description: string o source_id: int [ref: > DatabaseSourceMetadata.source_id]

C DatabaseSourceMetadata
<ul style="list-style-type: none"> o source_id: int [unique, primary key] o name: string o description: string o source_type: string o source_url: string o license: string o version: string o creation_date: datetime o update_date: datetime o keywords: string o publisher: string o contributor: string o language: string

C ManufacturingEquipment
<ul style="list-style-type: none"> o equipment_id: int [unique, primary key] o resource_id: int [ref: > DistributedResource.resource_id] o brand: string o model: string o type: string o description: string o source_id: int [ref: > DatabaseSourceMetadata.source_id]

C DataVerifier
<ul style="list-style-type: none"> o verifier_id: int [unique, primary key] o name: string o description: string o contact_id: int [ref: > Contact.contact_id] o address_id: int [ref: > Address.address_id]

C ManufacturingFacility
<ul style="list-style-type: none"> o facility_id: int [unique, primary key] o resource_id: int [ref: > DistributedResource.resource_id] o name: string o description: string o type: string o capacity: int o capabilities: string o source_id: int [ref: > DatabaseSourceMetadata.source_id]

C DistributedResource
<ul style="list-style-type: none"> o resource_id: int [unique, primary key] o address_id: int [ref: > Address.address_id] o contact_id: int [ref: > Contact.contact_id] o verification: bool o verifier_id: int [ref: > DataVerifier.verifier_id] o source_id: int [ref: > DatabaseSourceMetadata.source_id]



4.2. Data privacy policy

The map of machinery data privacy policy was needed for the data processing and collecting in compliance with GDPR regulations, the policy is found at:

<https://map.internetofproduction.org/data-privacy-policy/index.html> it covers the following regulations:

United States:

- California Consumer Privacy Act (CCPA) and California Privacy Rights Act (CPRA)
- Colorado Privacy Act (CPA)
- Connecticut Data Privacy Act (CTDPA)
- Utah Consumer Privacy Act (UCPA)
- Virginia Consumer Data Protection Act (VCDPA)

Europe:

- General Data Protection Regulation (GDPR), including the EU, UK, Iceland, Norway, and Liechtenstein.

Africa:

- Protection of Personal Information Act (POPIA) (South Africa)

The policy outlines the responsibilities on the handling of sensible data about the people residing in the covering regions. The relevancy of the United States covering is that of Data hosting services location, data sources and users of the platform that may request access through the web of the Map of Machinery data from the listed territories. The use of analytics also requires a policy in place.

Responsibilities in Data Aggregation

The policy defines key responsibilities to ensure data is handled ethically, securely, and transparently. In the event of a data breach, affected individuals will be promptly notified, and immediate containment measures will be implemented to minimize the breach's impact and prevent further exposure. Data owners have the right to request changes to their data or its deletion, ensuring transparency, user autonomy, and compliance with global privacy standards. The policy clearly defines the purpose of data collection and processing, along with any limitations, ensuring that users and stakeholders understand how their data will be used. A designated Data Protection Officer oversees compliance with privacy regulations and serves as a point of contact for privacy-related concerns. The policy also specifies the lifecycle of collected data, including how long it will be retained. For example, the retention period is measured in months to balance operational needs with privacy compliance.



4.3. ETL System

The ETL system was implemented in Python 3.12, using the library Metaflow, open-sourced by Netflix under Apache License Version 2.0. Metaflow serves as a pipeline orchestration wrapper, the code written as proof of concept for the Jupyter Notebooks was refactored to use the Metaflow pipeline description specification.

Metaflow is built around a few core concepts:

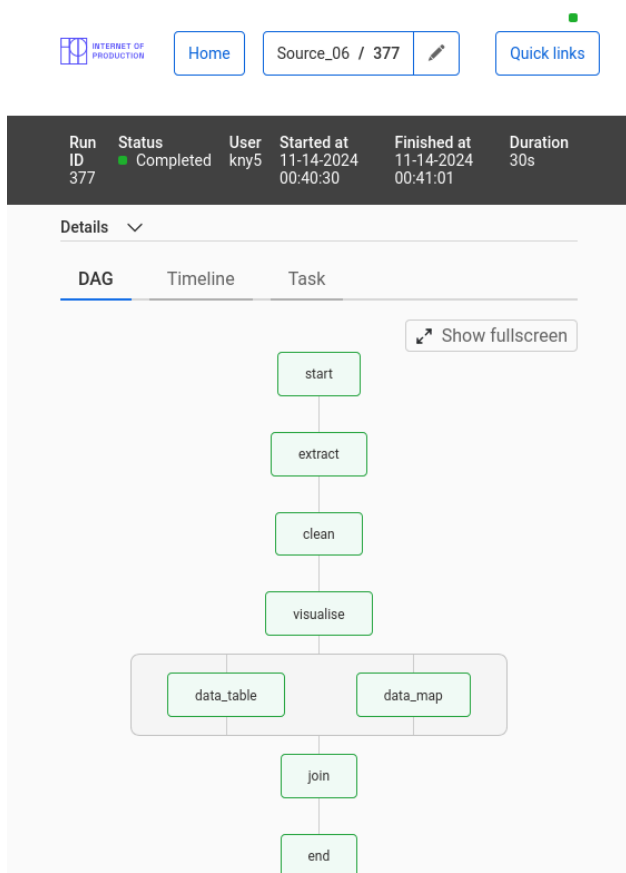
- **Flows:** A flow represents a collection of tasks that define the entire pipeline. It is implemented as a Python class inheriting from FlowSpec.
- **Steps:** Each method in a FlowSpec class, annotated with the `@step` decorator, represents an individual task within the flow. These tasks can be executed sequentially or in parallel.
- **Data Passing:** Metaflow allows easy passing of data between steps. Data is stored in class attributes, and the framework manages the movement of data between steps.
- **Parallelism and Scalability:** Metaflow offers built-in support for parallel execution and scaling across distributed environments using decorators like `@batch` for resource allocation and `@foreach` for parallel processing.
- **Versioning and Lineage:** Metaflow automatically tracks the history and lineage of data, ensuring reproducibility and easier debugging.

Main interface view of Metaflow showing the pipelines performed, some are tagged as failed, this was during the development process when the Python code was refactored to use the Metaflow Pipeline specification. Task Scripts were run on fail-safe for testing purposes.

Flow	ID	User	Started at	Finished at	Duration	User tags
JoinData01	258	kny5	11-05-2024 01:44:21	11-05-2024 01:46:13	1m 52s	development
JoinData01	257	kny5	11-05-2024 01:43:07	11-05-2024 01:43:07	0s	development
JoinData01	256	kny5	11-05-2024 00:57:38	11-05-2024 01:02:01	4m 23s	development
JoinData01	245	kny5	11-04-2024 12:42:24	11-04-2024 12:45:17	2m 52s	development
JoinData01	208	kny5	11-01-2024 18:38:55	11-01-2024 18:43:12	4m 16s	development
JoinData01	204	kny5	10-31-2024 13:15:50	10-31-2024 13:18:37	2m 47s	development
JoinData01	202	kny5	10-31-2024 12:00:36	10-31-2024 12:04:53	4m 16s	development
JoinData01	196	kny5	10-31-2024 11:22:36	10-31-2024 11:22:36	0s	development
JoinData01	179	kny5	10-31-2024 06:27:51	10-31-2024 06:30:37	2m 46s	development
JoinData01	178	kny5	10-31-2024 06:12:39	10-31-2024 06:15:55	3m 15s	development
JoinData01	177	kny5	10-31-2024 05:56:20	10-31-2024 05:59:05	2m 45s	development
JoinData01	176	kny5	10-31-2024 05:13:08	10-31-2024 05:15:56	2m 48s	development
JoinData01	175	kny5	10-31-2024 05:06:42	10-31-2024 05:10:00	3m 18s	development



Specialized software functions were coded to perform ETL Tasks for each Data Source extraction case and Pipeline. As follows:



Start:

Loads required libraries and establishes connections to PostgreSQL database, Artifact repository, and Server-side services.

Extract:

Performs data extraction through HTTP requests, database connections, or file downloads.

Stores raw data and creates a Pandas DataFrame artifact with the extracted and formatted values.

Clean:

Parses the extracted data into the desired structure (columns and rows). Applies filtering criteria to remove irrelevant or invalid entries. Produces a cleaned output DataFrame for further processing.

Visualize:

Processes the cleaned DataFrame, splits the process into: Rendering a tabular data representation, and an interactive Mercator projection web map.

Join and End:

Merges outputs from the parallel visualization tasks into a unified result Returns True if all steps are completed without errors. Flags the pipeline task as failed if any step encounters an issue.

Code Repository:


The ETL pipeline implementation is available in the Open-Know-Where Data Management Repository: https://github.com/iop-alliance/okw_data_management/tree/main/pipelines/metaflow Explore the code files for detailed insights into the pipeline structure and functions.

For more information of Metaflow please consult to the project documentation:

<https://docs.metaflow.org/introduction/what-is-metaflow>




Task details view of the ETL system using Metaflow. In this example, the pipeline extracts the data from source id 6 or Make Works. The system manages all the operations described in the pipeline flow process, stores all the information including failed runs and keeps a version control of the data generated. This ensures that even when the original data is not available, the tasks can take data from past successful runs and perform more data acquisition operations.



[Home](#)

Source_06 / 377 / start / 3623



Quick links

Run ID	Status	User	Started at	Finished at	Duration
377	● Completed	kny5	11-14-2024 00:40:30	11-14-2024 00:41:01	30s

Details ▼

DAG

Timeline

Task: 3623

▼ ▲

Artifact

Mode
Workflow ▼

Status
All (8) ● ● ● ▼

Order by
Started at ↓ ▼

Group by step

▼	Task	Duration	
▼	start (1)	1.0s	
	3623	0.1s	
▼	extract (1)	18.1s	
	3624	15.5s	
▼	clean (1)	1.3s	
	3625	0.5s	
▼	visualise (1)	0.9s	
	3626	0.1s	
▼	data_table (1)	1.4s	
	3627	0.5s	
▼	data_map (1)	1.3s	
	3628	0.5s	
▼	join (1)	1.1s	
	3629	0.2s	
▼	end (1)	1.0s	
	3630	0.1s	

Attempt 0

Task ID	Step	Status	Started at	Finished at	Duration
3623	start	● Completed	11-14-2024 00:40:31	11-14-2024 00:40:31	0.1s

Task details ▲

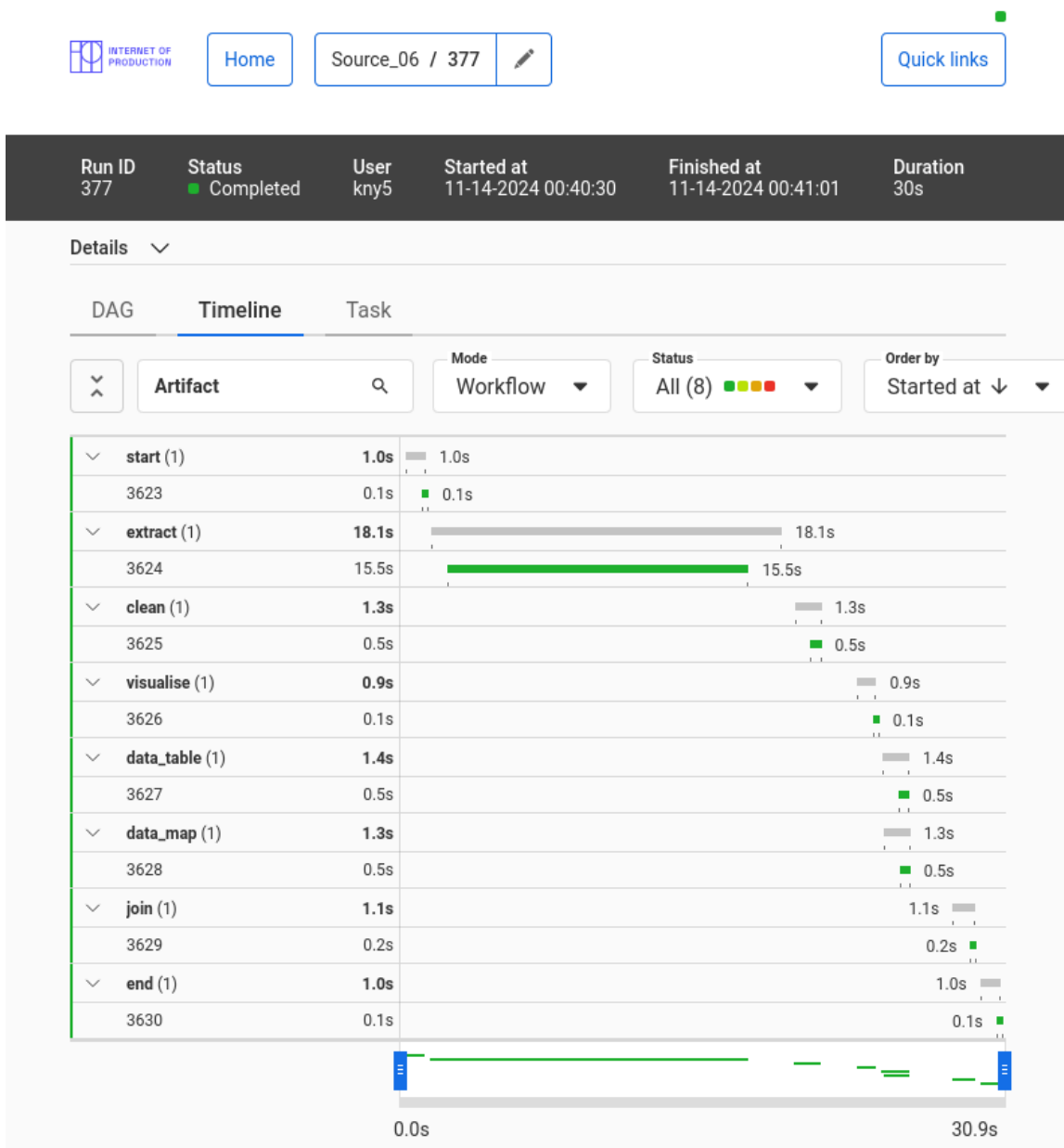
Metadata

attempt	0
attempt-done	0
attempt_ok (internal_attempt_status)	True
ds-root	s3://metaflow
ds-type	s3
metaflow_version	2.12.25
origin-run-id	None
python_version	3.12.5
runtime	dev
user	kny5

- Task info
- stdout
- stderr
- Artifacts



Timeline view of the ETL system: This graphical representation over time helps to understand the effectiveness of the operations performed, the extraction process in this example consists of multiple HTTP Get Requests timed one second apart one to each other to avoid server admin quota download violations.



For this example, the code is available at the repository:

https://github.com/iop-alliance/okw_data_management/blob/main/pipelines/metaflow/source_06.py



View of Tabular data representation of the Pipeline "Source06" Task "data_table":

Home
Source_06 / 377 / data_table / 3627

Quick links

▼	start (1)	1.0s
	3623	0.1s
▼	extract (1)	18.1s
	3624	15.5s
▼	clean (1)	1.3s
	3625	0.5s
▼	visualise (1)	0.9s
	3626	0.1s
▼	data_table (1)	1.4s
	3627	0.5s
▼	data_map (1)	1.3s
	3628	0.5s
▼	join (1)	1.1s
	3629	0.2s
▼	end (1)	1.0s
	3630	0.1s

Show 10 rows ▼
CSV
Excel
Search:

name	latitude	longitude	record_source_url
AnNead Hand Knitwear	56.95	-6.14	https://make.works/companie
Adam Plastics Ltd	55.74	-4.16	https://make.works/companie
Aero Leather Clothing	55.61	-2.8	https://make.works/companie
Alex Begg & Company	55.47	-4.62	https://make.works/companie
All Round Signs	56.85	-5.16	https://make.works/companie
Andrew Muirhead & Son	55.84	-4.22	https://make.works/companie
Araminta Campbell	55.970743	-3.180426	https://make.works/companie
Black Isle Bronze	57.58	-3.85	https://make.works/companie
Brodies Timber	56.56	-3.6	https://make.works/companie
Bute Fabrics Ltd	55.83	-5.06	https://make.works/companie

Showing 1 to 10 of 359 entries
« < 1 2 3 4 5 ... 36 > »

Task info

stdout

stderr

Artifacts

Card: html

View of same Pipeline with Task "data_map" rendering the map representation of the data:

Home
Source_06 / 377 / data_map / 3628

Quick links

▼	start (1)	1.0s
	3623	0.1s
▼	extract (1)	18.1s
	3624	15.5s
▼	clean (1)	1.3s
	3625	0.5s
▼	visualise (1)	0.9s
	3626	0.1s
▼	data_table (1)	1.4s
	3627	0.5s
▼	data_map (1)	1.3s
	3628	0.5s
▼	join (1)	1.1s
	3629	0.2s
▼	end (1)	1.0s
	3630	0.1s

Card: html
🗨️ 📄

Notice: This map and its data is a work in progress.

Task info

stdout

stderr

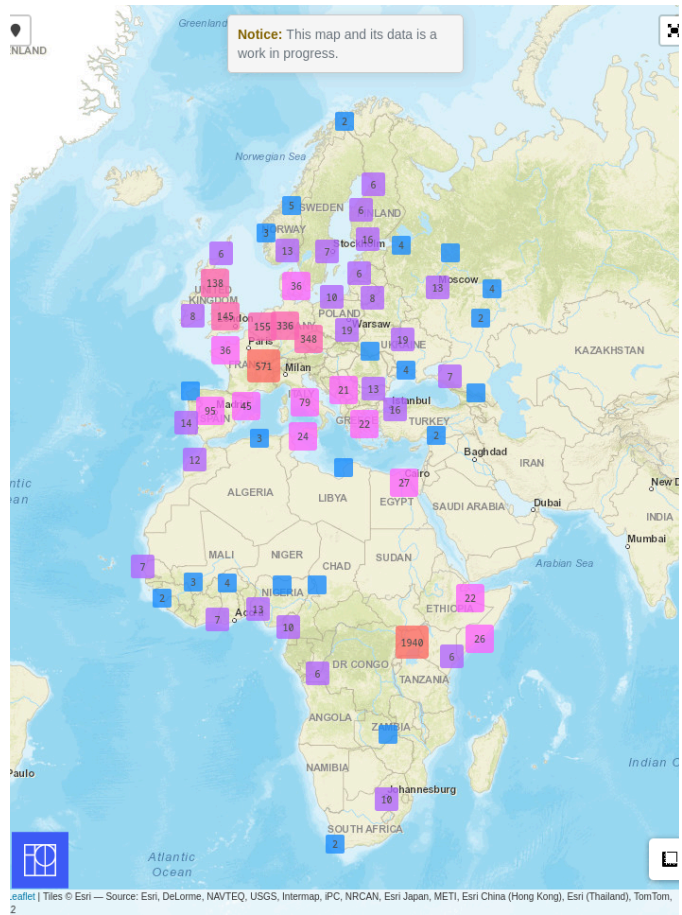
Artifacts

Card: html



5. Map design

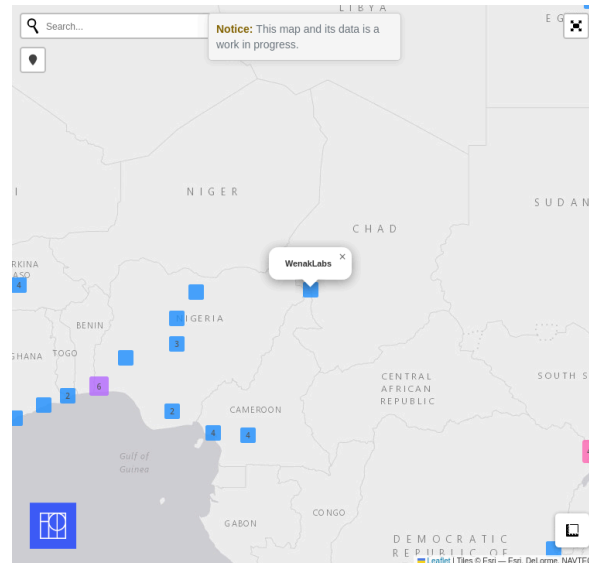
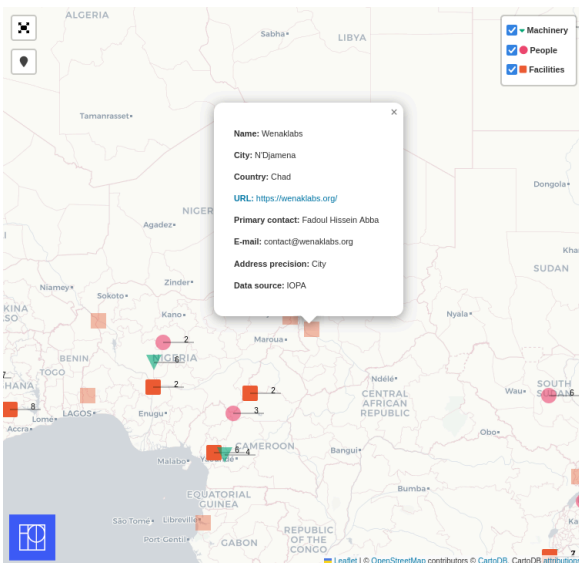
The visual design of the map had the mission to seek for differentiation and innovation. As there are other map systems already explored in the data assessment. One of the most notorious elements are the cluster markers programmed with a customized function that changes the size and colour of the squared markers depending on the quantity of facilities in the area covered. Other elements inside the map did not follow regulations and had to be removed after review of the GDPR regulations, the first version had a People’s data filter, it was removed and revisited. The new version includes different layers of data protection in place, preventing accessing data without consent.



Left image first design of map of machinery displaying unique location clusters, January 2024, total 556. Right image showing latest version unique locations 4372.

One of the key factors for the usability of the map of machinery is the rendering performance. The more markers, the more computational resources are needed for displaying the map. Keeping pruned data on the map can make a huge difference. The current size of the map with 11K markers is 483 KB. The first version shown at the left is 1.8MB with ~600 markers.





Map design comparison, left image first version holding all available information on PopUp tooltips. Right Image Only location name is shown, with a link on it redirecting to the data source. The size in the html file for the latest version is 100 times smaller in comparison.

Tools available on the map:

- **Near me search:** Zooms and locates the map to the current user's location. Permission for geolocation is needed.
- **Search bar for geolocation:** It searches and zooms over cities, countries or specific addresses.
- **Ruler:** A measuring ruler that measures linear Euclidian distance between two points.
- **Fullscreen:** Resizes the map to the full size of the current display used
- **Notification panel:** A floating panel that displays messages like disclaimers.

The clustering rules define the appearance of markers based on the number of child points in a cluster. For a single point, no text is displayed, and the marker colour is Dodger Blue. Clusters with 2 to 5 points are displayed in Light Blue. As the number of child points increases, the marker size and colour change to indicate density: Light Purple for 5–20 points, Medium Pink for 20–50 points, Hot Pink for 50–100 points, and Coral Red for 100–300 points. Clusters with 300–500 points are represented in Bright Red, and this colour continues for larger clusters with 500–2000 points, with marker size increasing to 40 for clusters exceeding 2000 points, which are displayed in Firebrick Red to signify the highest density.

The code used for the map design, Python for control and JavaScript for visualisation:

Python:

https://github.com/iop-alliance/okw_data_management/blob/main/pipelines/metaflow/_visualisations_.py

JavaScript:

https://github.com/iop-alliance/okw_data_management/blob/main/pipelines/metaflow/assets/cluster_icons.js



6. Minimal Viable Product: A map for disaster response

During the designing process a series of interviews were conducted within the mAKE consortium community members. The findings are documented here:

<https://community.internetofproduction.org/t/okw-working-group-action-plan-step-1-one-to-one-conversations/647>. The interviewing process aimed to gather relevant use cases to enhance understanding of the purpose and value of the Map of Machinery, ensuring it benefits communities and creates a meaningful impact.

The interview process followed after the request for writing a project charter, and finding clear projects to develop community engagement, these projects will serve to motivate community members to share data, provide feedback with a clear purpose and tangible outcome.



Image of Brainstorm of ideas, and needs for data and the map of machinery. Obtained during interview with consortium members.

Four use cases were identified:

- Disaster response,
- Economical development data,
- Learning for makerspaces,
- Data tools for hardware designers.

The disaster response scenario opened the interest in makers to share their information in an organic way.

Detailed use cases detailed can be found here:

<https://map.internetofproduction.org/use-cases/index.html>

Three more tasks were added for development:



- A data privacy protocol for disaster response.
- A tool for data visualisation is more versatile. Where data is loaded freely for data crossing.
- A guide for the usage.

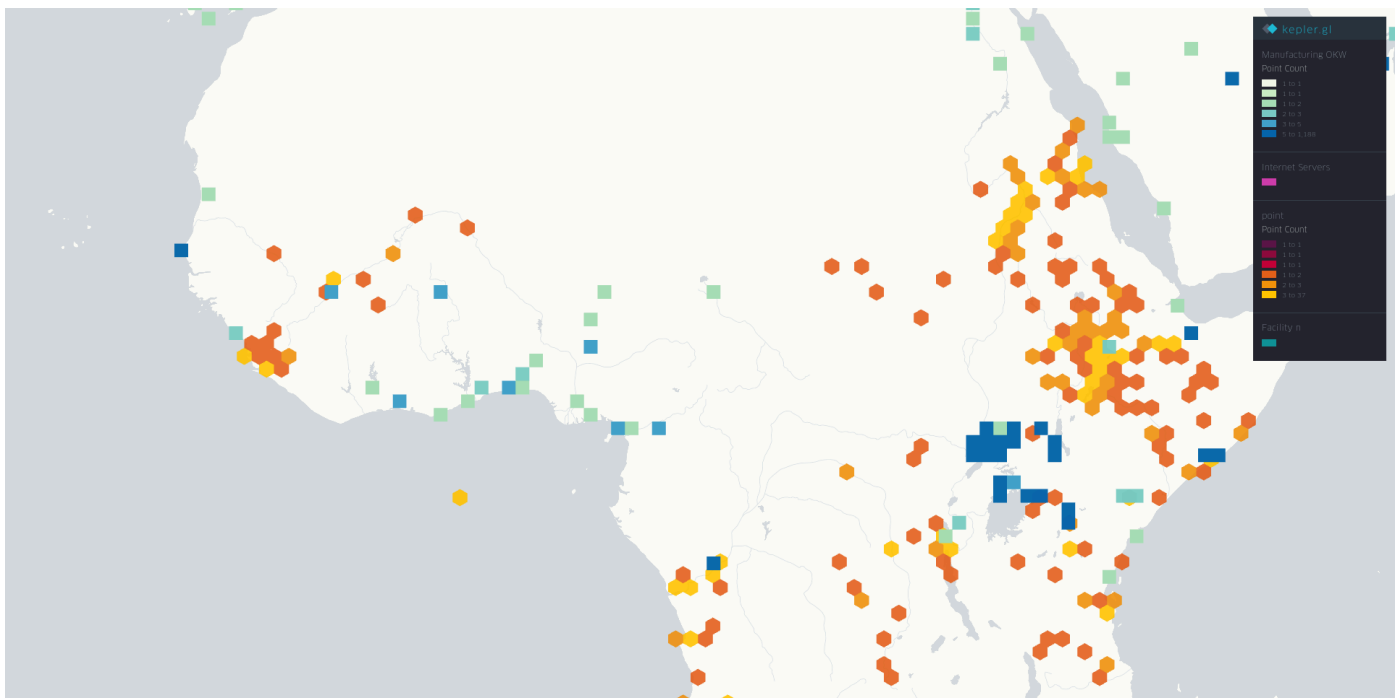
Protocol draft:

The data privacy override in case of disaster has the objective to open access to geolocation information in case of an emergency response, for those locations that have their location masked for privacy reasons. The protocol would enable temporary and limited access based on a range of current locations.

Data visualisation tool:

KeplerGL uses the data provided by the Map of machinery, and using the data addition tool, it's possible to cross-reference relevant data like in example: Public Hospitals Hexagons and makerspaces squares in Sub-Saharan Africa. Living map available here:

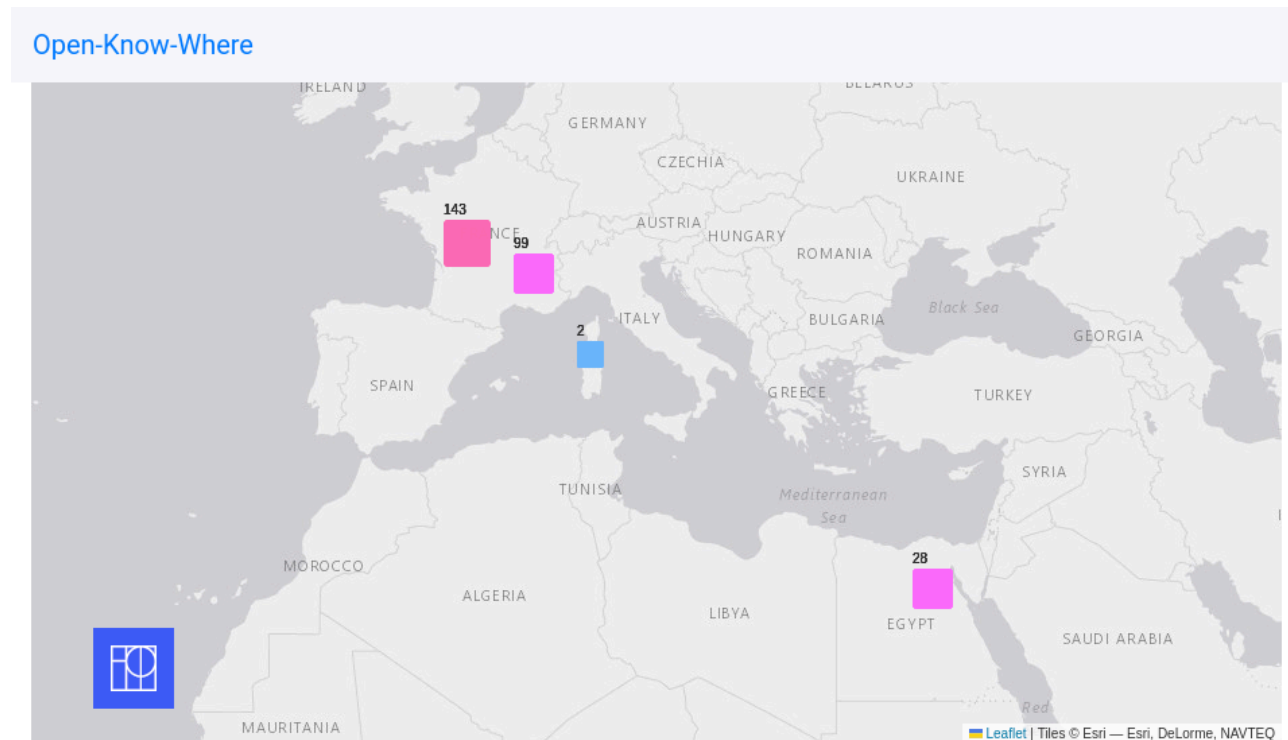
https://map.internetofproduction.org/hospitals_makerspaces_subsahara.html



Usage and MVP:

A Web Page that displays all Locations and machines available in a selected country under an emergency disaster situation together with the data of a counterpart in Africa or Europe for cooperation. In the example it shows Egypt and France. In the future the parameters will add the second option based on Geopolitical variables.

A web page for the displaying of this example is available at:
https://map.internetofproduction.org/disaster_response_mvp.html



Facilities

Show entries

Search:

name	city	email	capabilities
8 FabLab Drôme	Crest	contact@8fablab.fr	['three_d_printing', 'cnc_milling', 'circuit_production', 'laser', 'precision_milling', 'vinyl_cutting']
ACoLab	Clermont-Ferrand	contact@acolab.fr	['three_d_printing', 'cnc_milling', 'circuit_production', 'laser', 'precision_milling', 'vinyl_cutting']
AgriLab	Beauvais	agrilab@unilasalle.fr	['three_d_printing', 'cnc_milling', 'circuit_production', 'laser', 'precision_milling', 'vinyl_cutting']



7. Data collection Platform

7.1. Surveying tools

Tool	Advantages	Disadvantages	GPS Data Collection
Google Forms	<ul style="list-style-type: none"> - Easy to use with no setup required. 	<ul style="list-style-type: none"> - Limited advanced features for offline use or complex data collection. 	<ul style="list-style-type: none"> - No native GPS data collection functionality.
	<ul style="list-style-type: none"> - Integration with Google Sheets for data management. 	<ul style="list-style-type: none"> - Limited customization for survey logic and design. 	<ul style="list-style-type: none"> - Requires third-party integrations or manual input of location coordinates.
	<ul style="list-style-type: none"> - Free for basic usage. 	<ul style="list-style-type: none"> - Requires a stable internet connection; no native offline mode. 	
	<ul style="list-style-type: none"> - Accessible to non-technical users. 	<ul style="list-style-type: none"> - No built-in data privacy features beyond Google's policies. 	
KoBoToolbox	<ul style="list-style-type: none"> - Designed for humanitarian and development work, ideal for complex surveys. 	<ul style="list-style-type: none"> - Steeper learning curve for non-technical users. 	<ul style="list-style-type: none"> - Supports GPS data collection natively through mobile devices.
	<ul style="list-style-type: none"> - Supports offline data collection and sync. 	<ul style="list-style-type: none"> - Limited integration with third-party tools. 	<ul style="list-style-type: none"> - Can record latitude, longitude, altitude, and accuracy directly into forms.
	<ul style="list-style-type: none"> - Robust data analysis and visualization tools. 	<ul style="list-style-type: none"> - Hosted options may require server setup for advanced use cases. 	
	<ul style="list-style-type: none"> - Open-source and free version is available for humanitarian and research organisations under the UNHCR server. 	<ul style="list-style-type: none"> - Interface may feel less intuitive compared to other tools. 	
ODK	<ul style="list-style-type: none"> - Highly customizable and versatile for complex data collection scenarios. 	<ul style="list-style-type: none"> - Requires technical expertise to set up and manage. 	<ul style="list-style-type: none"> - Comprehensive GPS data collection functionality, including offline capability.
	<ul style="list-style-type: none"> - Strong support for offline data collection. 	<ul style="list-style-type: none"> - No native integration with widely used data tools like Google Sheets or Airtable without custom solutions. 	<ul style="list-style-type: none"> - Allows integration of GPS with additional sensors or metadata (e.g., timestamps, device info).
	<ul style="list-style-type: none"> - Open-source and free for most use cases. 	<ul style="list-style-type: none"> - Can be resource-intensive to host and maintain for large-scale deployments. 	
	<ul style="list-style-type: none"> - Supports multimedia data collection (e.g., images, GPS). 	<ul style="list-style-type: none"> - Less intuitive for non-technical users compared to Google Forms or Airtable. 	
Airtable	<ul style="list-style-type: none"> - Combines database functionality with a user-friendly spreadsheet interface. 	<ul style="list-style-type: none"> - Requires a subscription for advanced features and larger datasets. 	<ul style="list-style-type: none"> - No native GPS data collection functionality.
	<ul style="list-style-type: none"> - Supports collaboration and workflow automation. 	<ul style="list-style-type: none"> - No robust offline data collection features. 	<ul style="list-style-type: none"> - Requires third-party integrations or manual input for location data.
	<ul style="list-style-type: none"> - Provides templates and integrations with various tools (e.g., Zapier). 	<ul style="list-style-type: none"> - Limited survey design capabilities compared to dedicated tools like KoBoToolbox or ODK. 	



	- Intuitive for users familiar with spreadsheets.	- Data privacy and security concerns for sensitive information.	
--	---	---	--

A deployment of the data collection form was available for distribution using ODK, Airtable, Google-Forms, and KoboToolBox. With the Kobo giving the best data validation robustness, its open-source licensing and free hosting by UNHCR made it the best suitable tool for the purpose for the Map of Machinery.

7.2. Data collection form design

Following the General Data Privacy Regulation for the European Union, an explanation of the data processing, purpose and contact information is required, for an informed data contribution keeping the rights of the contributor for deletion, modification and privacy. Form available at: <https://map.internetofproduction.org/data-submission/index.html> is then divided in 3 sections

KoboToolbox Choose Language English

Internet of Production Data Collection for Open-Know-Where

Africa Europe Maker Innovation Ecosystem

What: This is a self submission tool for the OKW MAP project to survey and collect data on machines, facilities, people, and materials where distributed manufacturing happens.

Who: This is intended to be answered by Makerspaces, FabLabs, and innovation hubs within the Alliance and its projects.

Purpose: The Internet of Production Alliance collects the data to form a database of machines, facilities, people, and materials by location, to be analyzed and mapped for the Open-Know-Where initiative.

Data Processing: The data collected is processed only by the DevOps office.

Data expiration: 36 months from the date of publicly available data, expected on 2024/01 TBA.

Our data policy is found at https://map.internetofproduction.org/okwmapdata_policy.html

For any map related concerns and requests contact by email to: antonio.anaya@internetofproduction.org

***Participation agreement**
I agree to share my information for the purposes stated in the form description.

OK

Header, Information, and participation agreement

What: This is a self-submission tool for the OKW MAP project to survey and collect data on machines, facilities, people, and materials where distributed manufacturing happens.

Who: This is intended to be answered by Makerspaces, FabLabs, and innovation hubs within the Alliance and its projects.

Purpose: The Internet of Production Alliance collects the data to form a database of machines, facilities, people, and materials by location, to be analyzed and mapped for the Open-Know-Where initiative.

Data Processing: The data collected is processed only by the DevOps office.

Data expiration: 36 months from the date of publicly available data, expected on 2024/01 TBA.

Our data policy is found at

<https://map.internetofproduction.org/data-privacy-policy/index.html>

For any map related concerns and requests contact by email to: antonio.anaya@internetofproduction.org
Participation agreement* I agree to share my information for the purposes stated in the form description.



Facility, information, and geolocation information capture

* Name of Facility

Country where the facility is located.

URL

Webpage or social profile of the Facility.

* Country

City, Village, or locality where the facility is located.

* City or Locality

* Street name

House, apartment or building exterior number

* Postal Code

* Capture GPS location

latitude (x.y °)

longitude (x.y °)

altitude (m)

accuracy (m)



© OpenStreetMap & Yohan Boniface & Humanitarian OpenStreetMap Team | [Terms](#)

This section captures the minimum required data for identification and finding of a manufacturing facility, details of the data selection are detailed in the Data Architecture section.

The GPS capturing section collects metadata of the device used, such as the quantity of satellites, altitude and accuracy.

The collector needs to write the country, city and post code, as GPS collection is not exempt from failures. This information can be used for cross-checking and validation in later stages.

A Street name is also needed for the cases where GPS has accuracy greater than 15 meters.

House apartment and exterior number are not forcefully required as most locations won't require that detailed information for being located. Is still available for the cases where a big mall or building has other similar facilities, under the same address and street.

Contact information

Contact information agreement
I agree to share my information for the purposes stated in the form description.

OK

▼ **Contact information**

E-Mail
E-Mail of Facility or Person in charge.

Name of contact
E-Mail of Facility or Person in charge.

Phone

This section asks for a second acceptance agreement for sharing contact information.

E-mail, name of contact and phone number are all optional fields, for the African context we have found in later stages a phone number is most likely shared compared to an Email account for business inquiries.

Machinery inventory information

For the capability and machinery collection the form adds a capability only if a machine is captured. In this case the form has a field specialized in three types of digital machinery: 3D printer, laser cutter and CNC.

The form is using the Catalogue of Machines for input completion, one of the products of the data analysis performed in the early stages of the project.

For more information on the Catalogue refer to: https://github.com/iop-alliance/okw_machines_catalog

Finally after collecting at least one machine, the form is validated and sent to the server.

KoboToolBox offers the offline-collection mode, it requires the collector to download the Kobo Mobile app available for Android devices. For more details of the functionalities of Kobo for Android refer to KoboToolBox Documentation and the following blog entry:

<https://community.internetofproduction.org/t/okw-data-awards-round-2-toolkits-and-requirements/302>

▼ **Inventory of machines**

» **Select category**

Select a machine from the catalogue
Use the search to find Brand and/or machine models

3D Printing
 Laser cutting
 CNC

* **Select 3D Printer Machine**

FlashForge Creator Pro

Uncatalogued Machine?
I agree to share my information for the purposes stated in the form description.

OK

Quantity

-

+

Save Draft ⓘ



7.3. Evaluate data collection process

The following is a list of surveying exercises using the form designed using KoboToolBox, the data specification by Open-Know-Where for the Map of Machinery. For the Data Awards program funded by the Internet of Production Alliance:

- Mapping manufacturing capabilities in Uganda, by OpenStreetMaps Uganda:
 - Report: https://map.internetofproduction.org/awards_report_uganda.pdf
 - Locations mapped: 344.
 - Timeframe: 1 month.
- Mapping manufacturing capabilities in Nigeria, by Education4All Initiative.
 - Report: https://map.internetofproduction.org/awards_report_nigeria.pdf
 - Locations mapped: 50.
 - Timeframe: 1 month.
- Mapping manufacturing capabilities in Cameroon, by MBoaLab:
 - Report: https://map.internetofproduction.org/awards_report_cameroon.pdf
 - Locations mapped: 20.
 - Timeframe: 2 months.
- Mapping manufacturing capabilities in Greece, by BlueGreece:
 - Report: https://map.internetofproduction.org/awards_report_greece.pdf
 - Locations mapped: 15.
 - Timeframe: 2 months.
- Field Ready and NeedsList:
 - Report: https://map.internetofproduction.org/fr_report_africa.pdf
 - Locations mapped: 5766.
 - Timeframe: 6 months.

Takeaways

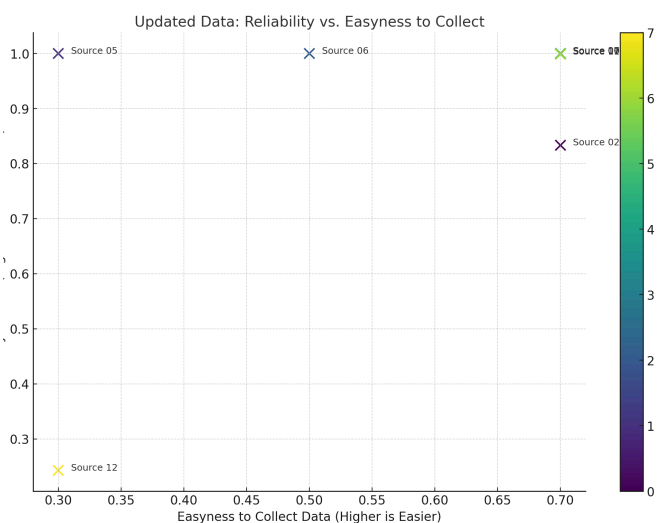
The surveying required for the African locations, the participation of several volunteers, hired by the local funded organisations, and the amount of data collected was dependable on the technical expertise, the area covering, the timeframe and the volunteer's digital literacy.

For some organisations, the work of data collection was one of their recurrent services. Their experience was collected during monthly meetings with the Internet of Production Communities.

Surveying in person, door by door had the most participation rates, web/online distribution of the survey has shown to be a slower process. People have concerns about the information shared and most will trust people they know in their communities or neighbouring ones.

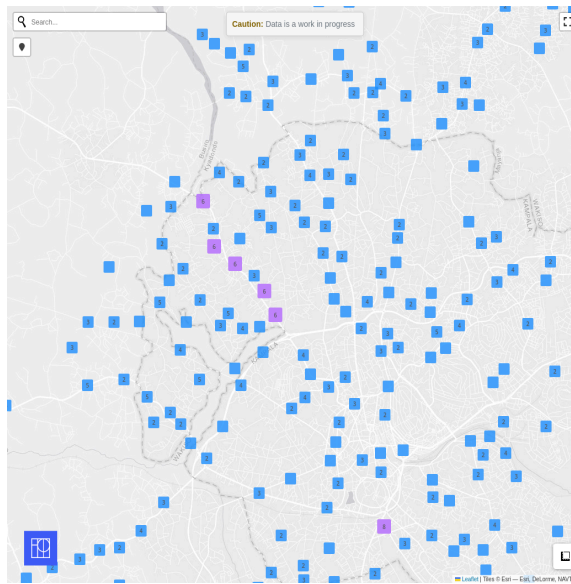
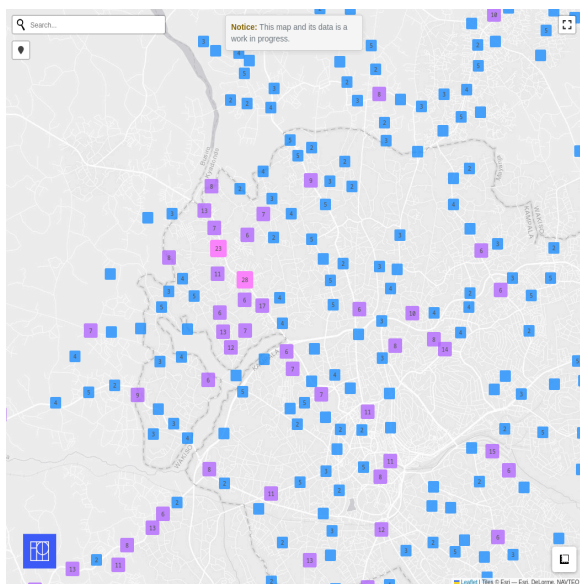


Data validation poses significant challenges, as illustrated by the COVID-19 Makers Response mapping efforts in Uganda and Nigeria. The observed data contained numerous single-point registrations, which introduced noise and duplicates. Since this data was collected externally and later provided to us, we were not directly involved in the collection process. The motivations for data collection varied; in some cases, volunteers were paid by an international NGO to document machinery locations and inventory data. This payment model may have inadvertently encouraged repeated submissions of the same location, further contributing to duplication and inconsistencies in the dataset.



KoboToolbox has proven to be an effective tool for collecting device and location metadata, enabling efficient data filtering.

For example, in the graph below, purple markers indicate areas of higher density compared to blue markers, illustrating the effectiveness of the filtering process applied in Source 12, which relies on KoboToolbox for surveying and data collection. The graph also highlights the reliability of network-driven surveying methods, as these leverage regional reach to ensure more comprehensive and accurate data collection.



The filtering criteria applied a clustering process to markers within a 500-meter range. Locations were further clustered based on key collisions involving name similarity, both phonetically and grammatically. This approach allowed the merging of locations with the same name within the specified range and facilitated the detection of input errors. The code for the data validation function can be accessed here: https://github.com/iop-alliance/okw_data_management/blob/main/pipelines/metaflow/_functions_.py

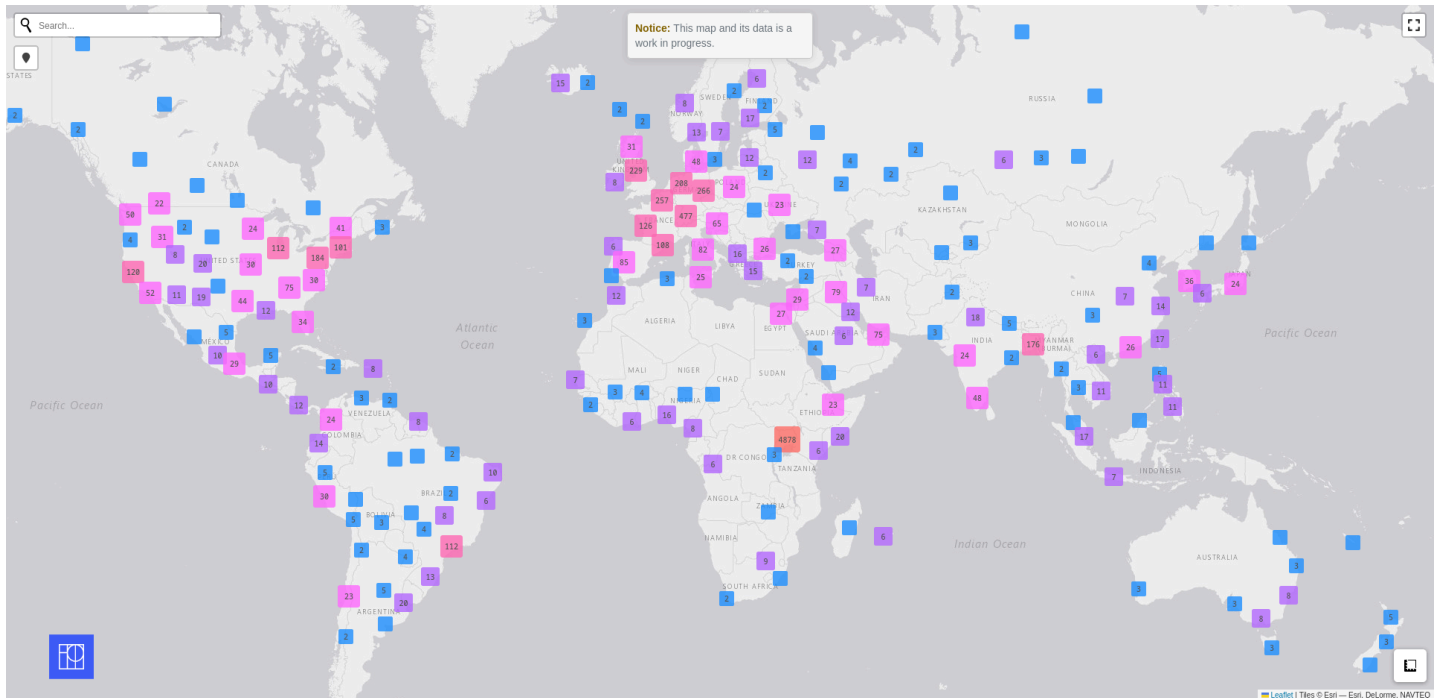


8. Data and map usage

8.1. Online Maps

Maps follow a visualisation method of clusters, choropleth Squares, that modify their colour and size, and number label based on the number of Geolocated facilities in the database.

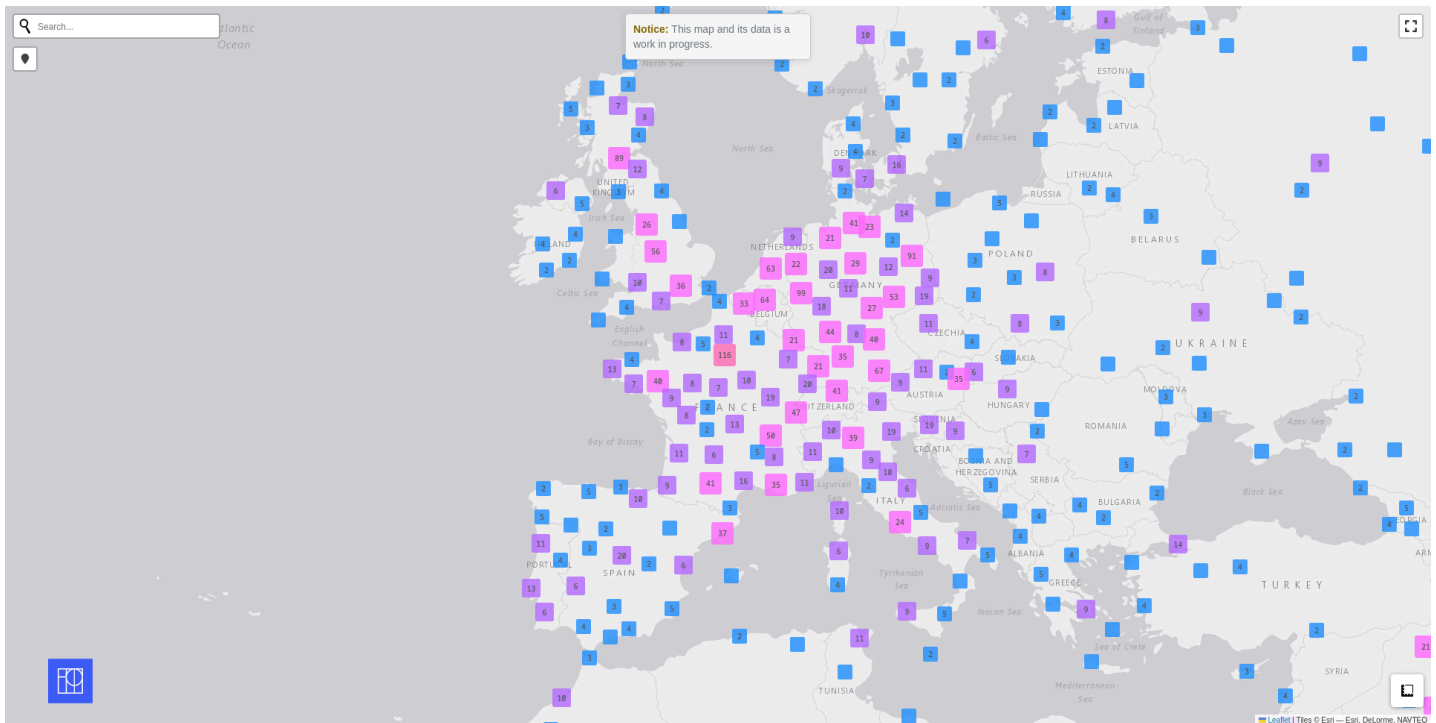
Interactive map available at: <https://map.internetofproduction.org/world.html>



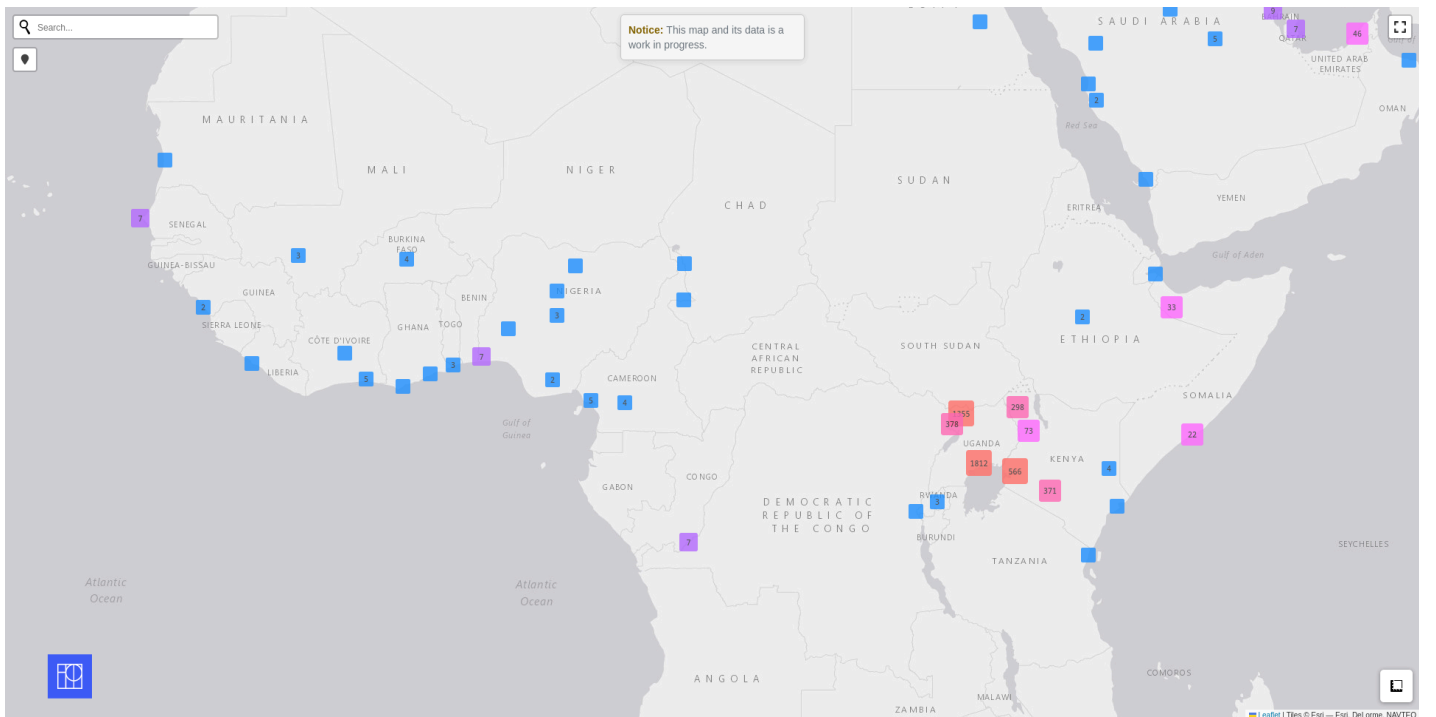
- 1) Worldwide data with the densest collected data is in Kenya, and Uganda. The reason is the data from the COVID-19 makers response data was aggregated. Raw data was provided by Field-Ready, and different filtering and validation methods are being used to output accurate information.



2) Europe: France, Germany and Italy show the most homogenic distribution of data. This is due to the aggregation of national network's data like Offener Werkstätten, Réseau Français des Fablabs and Make-Works network. Data here accounts more by self-submission.



3) Central Africa, Arabic Peninsula, data here is heavily dependent on surveying. And thus, the distribution regionally is scattered but denser in the case of West-Africa, due to NGOs work during the COVID-19 response.

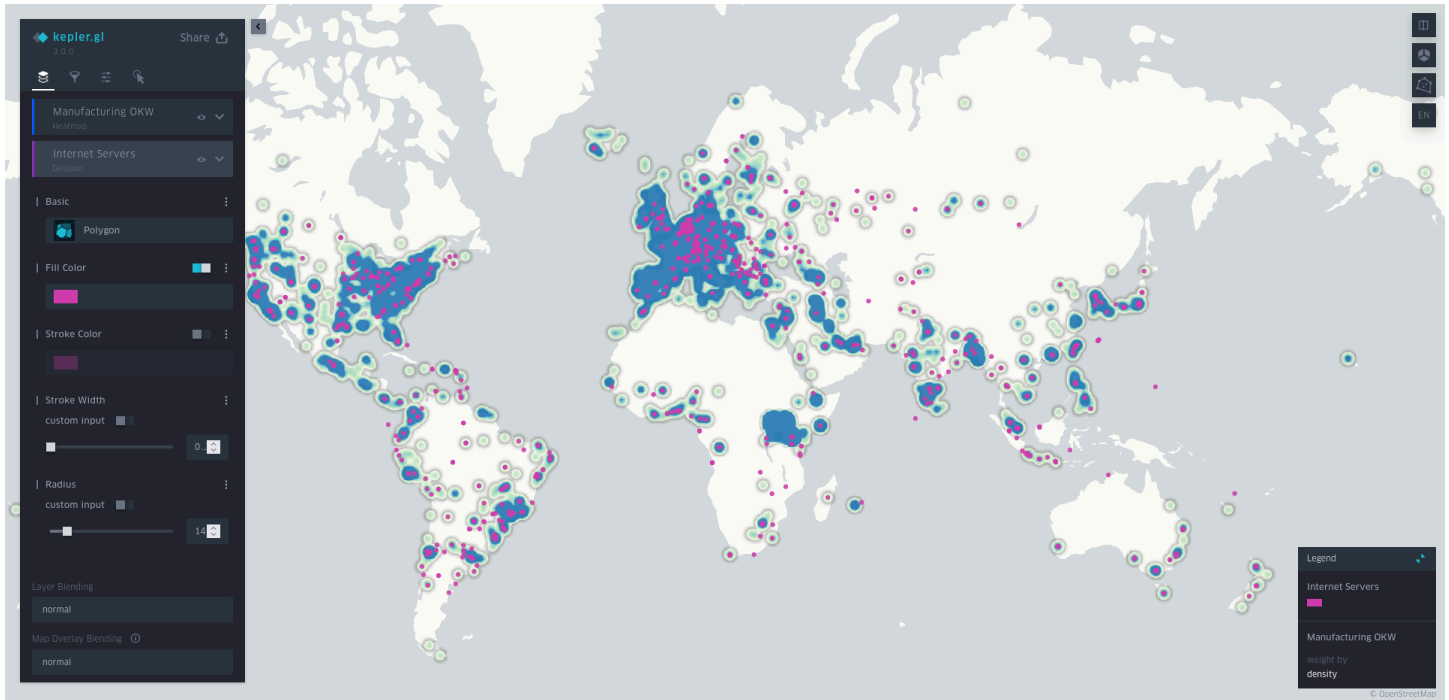


8.2. Data usage example

Internet availability and Map of Machinery data crossing map using KeplerGL. Data of Internet Exchange Points from IXBD EURO database.

With this data crossing we observe a relationship between internet providers infrastructure and the location of makerspaces.

The used is available at: <https://map.internetofproduction.org/locations.csv>

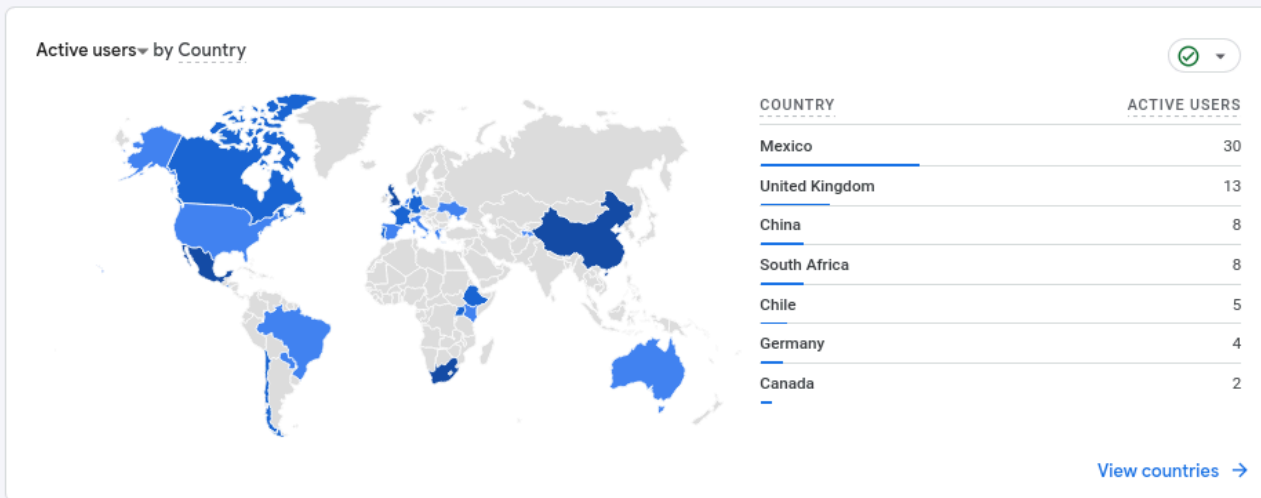


8.3. Data map analytics:

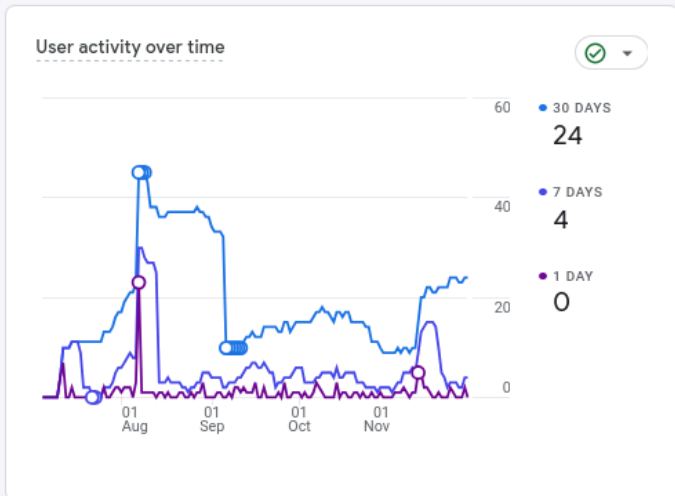
The analytics reveal significant user engagement trends linked to major events. The spike in activity in Mexico corresponds to the Fablab Conference in Puebla in August, which generated the highest number of active users. Similarly, increased engagement in South Africa aligns with the AfriLabs Conference in November, highlighting the platform's relevance in distributed manufacturing. The German Republica Festival in May also contributed to user adoption, reflected in active users from Germany. The user activity chart shows peaks in August and November, directly correlating with these events. New users primarily came through direct traffic and Google, indicating effective outreach during these conferences. The most viewed page, "Open-Know-Where Initiative," reflects the interest generated by showcasing the platform at these events. Key user interactions, such as "page_view" and "user_engagement," confirm meaningful engagement beyond casual visits. The map confirms geographic correlations, with high activity in Mexico, South Africa, and Germany, underscoring the impact of international events on platform adoption. Current date analytics available here: <https://map.internetofproduction.org/analytics/index.html>



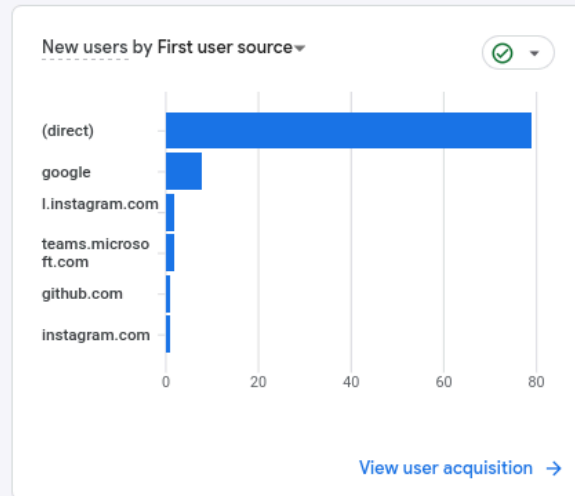
Analytics snapshot:



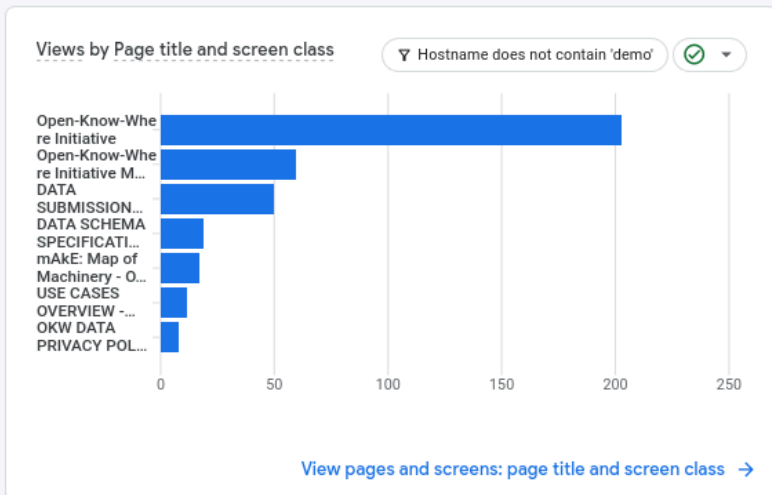
HOW ARE ACTIVE USERS TRENDING?



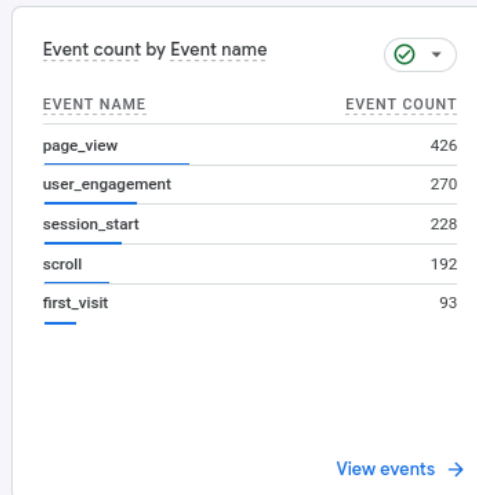
WHERE DO YOUR NEW USERS COME FROM?



WHICH PAGES AND SCREENS GET THE MOST VIEWS?



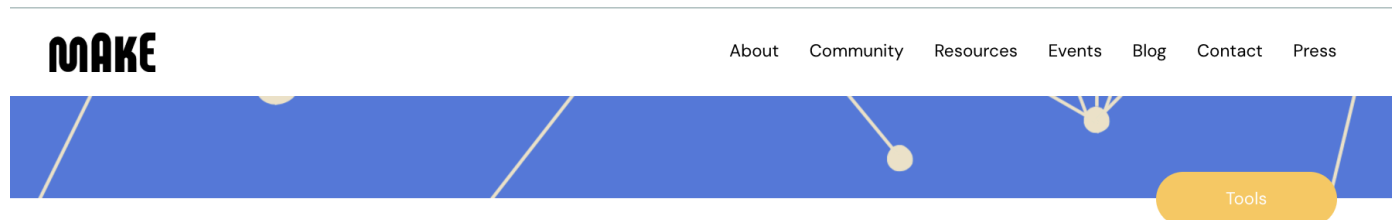
WHAT ARE YOUR TOP EVENTS?



9. Publication & Promotion

A series of promotional materials were created to promote and inform on the Map of Machinery, developed to provide data on the machines and spaces for makers, makerspaces and SMEs (small and medium-size enterprises) to be part of the distributed manufacturing network and participate in production.

- **A Blogpost**, shared on the [mAKE website](https://makeafricaeu.org/), presenting the origin of map of machinery, the use cases identified from discussions with the community, the collection of data from partners and global networks, how to add your space to the map and what comes next: <https://makeafricaeu.org/map-of-machinery/>



From Local to Global: The Power of the Map of Machinery

Are you a maker or SMEs looking for a specific machine? Are you starting your own project but you lack certain resources? Are you a company or a person who wants to manufacture something on a small scale? The Map of a Machinery is a tool for distributed manufacturing that will help you achieve your goals and explore the resources within the maker innovation ecosystem.

July 02, 2024

6 min read

- **A video** introducing the map and its missions and usage was also shared in the blogpost mentioned above, and on Youtube: https://www.youtube.com/watch?v=MJYyiH-xMnE&embeds_referring_euri=https%3A%2F%2Fmakeafricaeu.org%2F
- **Emails** were shared with the Consortium partners to share with their networks to collect more data to populate the map (English and French version).
- The map was presented at a series of **international events**, in-person and online:
 - Africa Makerspace Gathering (Nov. 2022), Cape Town – South Africa
 - re:publica – [Engaging communities, collecting data and solving problems](#) (May 2023), Berlin – Germany
 - GIG 10th anniversary – Workshop on the Map of Machinery, (May 2023) Berlin – Germany
 - [Fab24 Mexico](#) – Introduction to the Map of Machinery (Aug. 2024) Puebla – Mexico



- o [Rendez-vous Makers REFFAO](#) – A la découverte du project mAKE (Sep. 2024) online
- Leaflet presenting the technical innovative infrastructures developed
- Brochure on the Map shared with the Consortium partners networks
- Publication in the Africa Makerspace Magazine (Nov. 2024)
- Article submitted for publication in Distributed Design 3rd Edition, 2025

10. Learnings

10.1. Challenges

Keeping Participants Engaged

One of the recurring challenges is the quick disengagement of awardees and contributors after their initial involvement. This can leave datasets incomplete and disrupt the flow of valuable information. Building long-term engagement requires consistent communication, clear incentives, and ongoing support to keep participants motivated and involved.

Bad quality of the data

Data quality assurance is often overlooked during the collection phase, leading to issues like inaccuracies, duplicate records, or inconsistent data. Addressing these problems later in the process becomes time-consuming and costly. Implementing validation checks and quality controls from the outset is crucial to avoid these pitfalls.

Data Privacy concerns

Often people refuse to share their machinery information due to privacy and security concerns. The usage of GDPR and having a data schema architecture that enforces privacy does contribute to trust, but only under certain communities and scenarios. Digital literacy and skills are closely tied to the openness during data sharing. Financial benefits are important for short term impact of data sharing, although this also leads to duplicated and low quality data.

Technical challenges

Navigating existent solutions for the data management to identify the best possible combination of tools, design a transparent, reliable and maintainable framework and workflow, took most of the time during the development of the Map of Machinery. Half of the success of an open-source community project is the momentum buildup inside the global maker community, the other half is the personal investment to defy current paradigms and deadlines.



11. Technology Readiness Levels

Technology readiness levels (TRLs) are a method for estimating the [maturity](#) of technologies during the acquisition phase of a program. TRLs enable consistent and uniform discussions of technical maturity across different types of technology. TRLs are based on a scale from 1 to 9 with 9 being the most mature technology.

TRL	European Union	Online map of machinery
7	System prototype demonstration in operational environment	<p>Assessed in June 2022</p> <p>The Open know-Where standard was adopted and online prototype systems were used during the Covid 19 pandemic by Field Ready, NeedsList and HOT</p>
8	System complete and qualified	<p>Assessed at September 2022</p> <p>System complete and qualified Map.internetofproduction.org and API, backups etc Updates to Open Know-Where (unique IDs etc) Scraping of data Feedback from committee / technical experts on architecture</p>
9	Actual system proven in operational environment (competitive manufacturing in the case of key enabling technologies, or in space)	<p>Assessed at January 2023</p> <p>Use of more data from other sources Use of data for distributed manufacturing API being used to integrate with external services Feedback from users into system usability and functionality Expansion of the system to cover more 'resources' beyond the Open Know-Where standard (including, for example, how many of what items have been made where as an indicator of quality) No longer a prototype; part of the infrastructure of the Internet of Production. Open Know-Where standard is used by CosMYX 3D printers, and conversations with other machine manufacturers such as Mekanika. Integration into FabCityOS software</p>

Table 2. Technology Readiness Levels from 7 to 9



12. References

Open Know-Where Specification <https://standards.internetofproduction.org/pub/okw/release/2>

Open Know-Where Chapter <https://map.internetofproduction.org/okw-charter/index.html>

Open Know-Where Privacy Policy <https://map.internetofproduction.org/data-privacy-policy/index.html>

Open Know-Where Initiative <https://map.internetofproduction.org/index.html>

Data Submission Form

<https://map.internetofproduction.org/data-submission/index.html>

Open Know-Where Data Schema Specification to manage information about distributed resources

<https://map.internetofproduction.org/data-schema-design-choices/index.html>

Internet of Production Repository <https://github.com/iop-alliance>

Open Know-Where map of facilities and machinery - Data management repository

https://github.com/iop-alliance/okw_data_management

Open Know-Where machines catalog (WIP) https://github.com/iop-alliance/okw_machines_catalog

Open Know-Where Geonode scripts https://github.com/iop-alliance/okw_geonode_scripts

Open Know-Where Working Group <https://map.internetofproduction.org/working-group/index.html>

Open Know-Where Workshop: Exploring Use Cases for Open Know-Where

<https://docs.google.com/spreadsheets/d/16JdaPYkARnymJtSOKRGZ7vAhQIPsPRmvlhhr3i4XwtU/edit?usp=sharing>

Open Know-Where [2024 kickoff meeting: A conversation around Governance and new streamlined Action plans](https://community.internetofproduction.org/t/okw-2024-kickoff-meeting-a-conversation-around-governance-and-new-streamlined-action-plans/637/1)

<https://community.internetofproduction.org/t/okw-2024-kickoff-meeting-a-conversation-around-governance-and-new-streamlined-action-plans/637/1>

Open Know-Where Data Awards.

<https://www.internetofproduction.org/news/okw-manufacturing-world-map-data-awards-2>

Open-Know-Where data-schema. Considerations and design choices

<https://community.internetofproduction.org/t/okw-data-schema-definition-and-design-choices/662>



Open Know-Where miro board <https://miro.com/app/board/uXjVOawxGDO=/>

4 Use Cases

<https://miro.com/app/board/uXjVOawxGDO=?moveToWidget=3458764581834346574&cot=14>

Use Cases Overview <https://map.internetofproduction.org/use-cases/index.html>

mAKE: map of Machinery <https://map.internetofproduction.org/makeafricaeu/index.html> -
<https://map.internetofproduction.org/makeafricaeu.html>

Disaster response Open database

<https://map.internetofproduction.org/disaster-response-open-database/index.html>

Open Know-Where Working Group, Action plan step 1: one to one conversations

<https://community.internetofproduction.org/t/okw-working-group-action-plan-step-1-one-to-one-conversations/647>

Open Know-Where Working Group Rolling Notes

<https://community.internetofproduction.org/t/okw-working-group-action-plan-step-1-one-to-one-conversations/647>

Open Know-Where recap video <https://youtu.be/CCYOgct62ww?si=B64NOjTxRuJILFKT>

Fab24 proposals

<https://docs.google.com/document/d/1ukcLWjf7M00VF8FPXXmYPT-luUD9gYNr/edit?usp=sharing&oid=118248051350828216234&rtpof=true&sd=true>

Talk at re:publica24 – Engaging Communities, collecting data and solving problems with Antonio Anaya

<https://re-publica.com/en/node/5392>

Fab24 Mexico – Workshop

Local Manufacturing going global – Press release 2021

<https://www.internetofproduction.org/okwpressrelease>

Open Know-Where Launch recording <https://youtu.be/ZyTTwTxgO64?si=9UbOmgdleSPSD8GN>

Humanitarian Open Streetmap Team <https://www.hotosm.org/>

Technical Readiness Level (TRL) https://en.wikipedia.org/wiki/Technology_readiness_level



12.1. Works Cited

Andrew Lamb & Barbara Schack. (2021). From Distributed Design to Participatory Production. In D. Design, *This is Distributed Design* (pp. 220–230). Under Creative Commons Attribution–NonCommercial–ShareAlike 4.0 International License.

Tomas Diez from Fab Lab Barcelona at IAAC and Peter Baeck from Nesta (2020). Viral Design. in D. Design, Introduction, the DIY and Open Hardware Response to the COVID–19 Crisis (pp. 14–21). Under Creative Commons Attribution–NonCommercial–ShareAlike 4.0 International License.

